STANDARD TESTS

RUSSELL

REFERENCE LIBRARY

OF THE

STATE TEACHERS COLLEGE

WHITEWATER, WISCONSIN

No. 27722

Bought 4-30 *19*31.

Students of the State Teachers College are allowed to draw the books for use at their rooms on the following conditions:

1. No book must be taken from the library until it has been properly charged to the student drawing it.
2. Bound magazines, most Reference books and Reserved books may be taken from the library at 4. p. m. and must be returned at 8. a. m. the following day.
3. All other books may be kept for two weeks.

STANDARD TESTS

A HANDBOOK FOR THE CLASSROOM TEACHER

BY

CHARLES RUSSELL, PH. D.
PRINCIPAL OF THE MASSACHUSETTS STATE NORMAL SCHOOL AT WESTFIELD

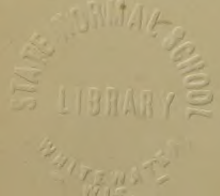
LB
1131
R87

GINN AND COMPANY
BOSTON · NEW YORK · CHICAGO · LONDON
ATLANTA · DALLAS · COLUMBUS · SAN FRANCISCO

COPYRIGHT, 1930, BY CHARLES RUSSELL
ALL RIGHTS RESERVED

PRINTED IN THE UNITED STATES OF AMERICA

630.1



The Athenæum Press
GINN AND COMPANY · PRO-
PRIETORS · BOSTON · U.S.A.

PREFACE

Standard Tests has been written for the teacher in service and for the student of teaching. It is intended to be a handbook neither so brief as to be difficult nor so extensive as to exceed that from which a good teacher might reasonably profit. It is so arranged as to provide for an understanding of tests and testing and is so planned that technics and processes may be examined apart from and together with their specific uses.

A consideration maintained throughout the book, and, it is to be hoped, throughout all the professional activities of the writer, has been the need and problem of the classroom teacher. The pupil to be taught — and, through him, the teacher who is to direct that teaching — is the primary focus of attention. Whatever will aid that teaching — whatever will make it more effective and whatever will result in a better or more complete understanding of the process by which it takes place — is worthy of the consideration of the student who would become a teacher or of the teacher who would become a better one. It is in this light and for this purpose that *Standard Tests* has been written, and it is in these terms that it should be read.

To those who have wittingly or unwittingly contributed to the writing the author gratefully acknowledges his indebtedness. Among the many stand preëminently Dr. William A. McCall and Dr. David Eugene Smith. The influence of the former may be found revealed upon every page as clearly as it is in the lives of his students; the influence of the latter, though less tangible, has been to provide by

illustrious example an enduring faith in teaching and in the classroom teacher.

To the students who have contributed to the thought of this volume, to the classroom teachers whose problems the writer has been privileged to consider, to Miss Marie Lerche, who has provided much material as well as aid, and to Miss Ida R. Abrams, who by patient consideration and untiring effort has helped to bring the work into being, the writer is deeply grateful.

CHARLES RUSSELL

CONTENTS

PART I. TESTING

CHAPTER		PAGE
I.	Measurement	3
II.	The Background of School Measurement	14
III.	The Development of School Tests	29

PART II. TESTS

IV.	Types of Tests	56
V.	The Construction of Tests	93
VI.	The Giving and Scoring of Tests	119

PART III. MEASURES

VII.	Measures of Mass	143
VIII.	Measures of Points	168
IX.	Measures of Achievement	201
X.	Measures of Ability	248
XI.	Measures of Rate	267
XII.	Measures of Effort	292

PART IV. USES

XIII.	Classification of Pupils within a School	304
XIV.	The Diagnosis of Individuals	368
XV.	The Diagnosis of Classes	410
XVI.	The Diagnosis of Schools	430
XVII.	The Rating of Pupils	453
XVIII.	Testing in Teaching	474
INDEX		511

STANDARD TESTS

PART I. TESTING

CHAPTER I

MEASUREMENT

Measurement and knowledge. Measurement is necessary alike to the acquisition, extension, or utilization of knowledge. It is a precision instrument necessary in research, an exact tool necessary to experimentation, and a means by which knowledge can be made available for use.

Progress in the development of knowledge is dependent, it is true, upon the insight, resourcefulness, inventiveness, and ingenuity of man, just as it is also dependent upon the extent and quality of previously acquired knowledge; but all these qualities in man that permit of progress in knowledge are also dependent upon the measures with which they are able to deal. It is remarkable, for example, that Galileo, with the crude instruments for the measurement of time and distance at his command, was able to accomplish the scientific advances which he made; it might well be questioned what his native resource, his keen insight, and his tremendous ingenuity might lead to today with the exact measures that he could now use.

For many centuries man was handicapped in his pursuit of knowledge by the crudity of his measurement units as well as by the crudity of his measuring instruments. However, when the measurement units themselves were refined, great results were possible even with crude instruments of measurement.

Such was the degree of accuracy secured by these early surveyors that Petrie found the maximum error in fixing the length of the sides of the Great Pyramid to be only 0.63 of an inch, or less than $\frac{1}{14\,000}$ of the total length, and the angle error at the corners to be 12″, or only $\frac{1}{27\,000}$ of a right angle.[1]

Early peoples, such as these Egyptians, developed certain measures to a high degree and were able to achieve astonishing results, both mechanically and scientifically, in the fields so developed. In other fields they were restricted. In surveying, in some of the mechanic arts, and in such astronomical observation as could be accomplished without the use of the telescope, these people were highly proficient. They were able to construct buildings with stonework so nicely fitted and so accurately measured as to leave us lost in wonder over their achievements. They were able so to survey the flow of their rivers that rechecks thousands of years later can find hardly a deviation. They were able to construct a calendar which in some respects was more useful than our own.[2]

Gains in some fields and losses in others, resulting in net gains in knowledge, culminated in the ascendancy of the Greeks as represented by Aristotle. Remarkable as this man was, his writings representing the height of knowledge up to his own time, nevertheless in the fields which today have been so clarified through measurement and measurement units some of his observations are absurd, not because of his lack of insight or his ability to observe but because of his inability to measure.

Through the centuries that followed, a time when general ignorance increased and specific knowledge declined in many fields, the progress of measurement was dormant. Ignorance bred superstition in the effort to explain natural phenomena;

[1] David Eugene Smith, *History of Mathematics*, Vol. I, p. 43. Ginn and Company, Boston, 1923.

[2] Ibid. chap. ii, sect. 5.

ignorance developed the astrologer rather than the astronomer; ignorance made the alchemist rather than the chemist, and the leech rather than the physician; ignorance resulted in the magician rather than the scientist. Ignorance today brings about the same result. Lack of knowledge, inaccurate knowledge, inadequate knowledge, all result in ignorance; but the human mind must work. Superstition plus speculation must result in guesswork. Guesswork plus inadequate measures of distance, weight, volume, and progress must, in its turn, result in quackery. Even as late as the beginning of the nineteenth century the plight of medicine was so great as to result in the death of Washington through bloodletting.[1]

Measurement and scientific progress. In one field of human knowledge after another progress has been made through the refinement of measures and the development of better means of measurement.

Early measures of distance were crude. Frequently the same name was given to measures of different lengths, as is true even today. The mile must be defined as the Roman mile or as the English mile, and so on, each mile having a different length. The length of the cubit, an ancient measure, is still in question, and the size of the Ark remains unknown.

Modern measures of distance and modern means of measuring distance have been increased in refinement as well as in accuracy. In many automobiles today the operation of the engines depends upon measurements as small as thousandths of an inch. Any garage mechanic must be able to set valve tappets accurately to within $\frac{2}{1000}$ of an inch, and the sleeve-valve motor runs efficiently with a clearance of no more than that. Vernier calipers will measure accurately to $\frac{1}{10000}$ of an inch when used even by those who are inexpert. By means of the microscope it is possible to measure the size of

[1] John Marshall, *Life of George Washington*, Vol. V, p. 761. C. P. Wayne, Philadelphia, 1807.

bacteria to $\frac{1}{10000}$ of a millimeter and to utilize such measures as a partial means in their identification.

In matters of weight tremendous advances have been made in the refinement of measures and in the precision of instruments. On the one hand, it has been possible to measure the mass of the great star Betelgeuse; on the other hand, there are scales which can distinguish differences in weight of $\frac{1}{50,000,000}$ of an ounce.

Whereas at one time speed was calculated in terms of days or minutes, it can now be calculated in infinitesimal parts of a second. There is the commercial speedometer for measuring speeds of ordinary travel, and there are also instruments for determing the speed of rays of light and the movement of molecules.

In temperature, however, where the ordinary thermometer can register, at most, two or three hundred degrees, it has been possible to measure to within 1 degree of absolute zero (− 272° C.) and as high as 14,000 degrees Fahrenheit.

In all the various scales of measurement which have been established for natural phenomena, such as distance, weight, volume, temperature, speed, and the like, the refinements in the measuring instruments, as well as in the measurement units themselves, have been the basis for the development of the sciences.

Chemistry without measurement was alchemy. With increased accuracy in measurement chemistry began to advance. First, the elements were isolated, and this led to the production of the Mendeléef table of the periodic elements. The gaps in this table, as studied by students in colleges of a few years ago, have largely been filled. Inert elements, — argon, neon, and the like, — whose presence was revealed by the gaps in the table, are now known and have been isolated.

Furthermore, accuracy of measurement has enabled extensive research into the character of matter. Electrons are

no longer mere theory, and their activities and place are now known. This knowledge of molecular construction has revolutionized some industries.

In physics, in the field of electricity alone, accuracy of measurement has within a decade permitted tremendous advances, and its measures have become household words. The boy of today can and does speak as intelligently of kilocycles, millifarads, ohms, and watts as the boy of a generation ago spoke of feet, yards, and inches.

In medicine the achievements in bacteriology, nutrition, physiology, X ray, and radium therapy have been made possible because of the refinements in measurement and measuring instruments.

Measurement and transportation. Through refinement in measurement the daily living of people has been changed, improved, and accelerated. Transportation has been increased by the development of railroads, motor cars, boats, highways, and airplanes.

Railroad trains run on steel rails, made possible by the measures for high degrees of heat, for the composition of metals, for the standardization of materials, for the detection of flaws, and the like. The trains run on the rails with safety because of signal devices that operate automatically. The trains run according to time schedules, which in the great terminals depend on accurate measurements of time in seconds. Almost every item in modern railroad transportation has been made possible through increased knowledge due to refinements in measurement in chemistry, in physics, and even in psychology.

Motor transportation has not only revolutionized industry but has itself been revolutionized by the demands made upon it. Highways in all parts of the country have been developed through the increased use of measurement. Motor fuels have been made available through the chemical

researches of the scientist, who has depended on the accuracy and the increasing refinement of his measurements. Safety in the use of automobiles has increased remarkably in the past decade because of the increased ability of manufacturers to measure the strength and character of the materials used in production, and because of the invention of new combinations of metals which are stronger, tougher, and lighter, and in which successful manufacture is dependent on accuracy in measurement of materials and temperatures.

Air transportation has had to depend for its success on the development of measures of the new medium which is now being utilized. The measurement of air currents and the development of altitude meters, pressure meters, and many other devices are the necessities for successful aviation. Until such measures have been developed to the point where they are unfailingly accurate, air transportation will remain in an experimental stage, although there has been much recent progress in this field.

Measurement and communication. The development of means of communication has been a long but rapidly accelerating process. Not until the invention of the printing press was the general means of communication much improved over that of the previous thousands of years. With that invention, which is dependent on the accurate measures of movable type, communication by means of the printed word expanded rapidly, and within a relatively few years books were available for all who could read.

After the printed word came the code word by means of the telegraph. This invention depended even more on measurement than did that of printing, for it involved the measurement of a new power — electricity. Telegraphy developed to parallel the railroads, and the safety of early railroad transportation depended largely on telegraphic communication along the lines.

The third development in communication, the telephone, was, like the telegraph, dependent on more exact knowledge and more exact measurement; but whereas the telegraph was dependent for its successful operation on experts familiar with sending and receiving messages by code, the telephone was an advance, in that it required no expertness in these two fields and thus could be more widely and usefully distributed.

A fourth development, dependent upon measures in other fields, was that of the phonograph. Here communication was not as spontaneous or as versatile as in the telephone and telegraph, but it was capable of wider distribution. The development of the phonograph has been the result of inventions and additions to human knowledge in many seemingly unrelated fields of science; for example, the increased effectiveness of the phonograph record is due largely to completer knowledge of the chemistry of halowaxes and to such products as redmanol and bakelite.

Probably one of the most effective means of communication that has yet appeared, combining, as it does, the international spread of books with the personal touch of the telephone or the phonograph, is the radio. Starting as "wireless," this was at first restricted because of the necessity for the use of code in sending and receiving. "Wireless" became radio with the development of sending apparatus capable of wider communication, and receiving apparatus capable of reproducing a wider variety of sounds. The importance of radio lies not so much in its advantage over the telephone or the phonograph, but rather (like science and the arts) in its indifference to national boundaries and provincial customs. For its successful operation it has been largely dependent upon the parallel development of adequate means for measuring and for reducing interference, for improving reception, for the making of tubes and appliances, and

for the development of power. Without these measures and the refinement of their accuracy, radio would be impossible.

Measurement and shelter and clothing. In all the processes developed in the preparation and manufacture of means of shelter and clothing, measurement has played a large part.

Spinning, weaving, knitting, and cutting have all been developed through refinements in measures and means of measurement in all the mechanical processes. Measurement has brought about the standardization of materials and manufactures, and this has stabilized prices and processes. Research in the field of chemistry applied to the development of aniline dyes has made possible such colors, and such permanency of colors, as were never known before. Shoes, boots, slippers, soles, rubber heels, rubbers, and the like (as they are now produced) have all been made possible through the refinement of measures and the standardization of processes.

In the field of shelter, outstanding developments due to the increasing refinement of measures have brought about comforts in living, now considered as necessities, which in a previous age would have been considered luxuries. Methods of building, building materials, household appliances, heat, light, power, and protection have all been developed through refinement in measures and means of measurement. Gas for cooking and electricity for lighting would be impossible without proper measures. Oil for heating, water for its manifold uses in the household, washing machinery, ironers, dishwashers, mixers, and appliances seemingly without end owe their use and their being to the development of adequate measures.

Measurement and food supply. There is probably no field where the development of adequate measures has so changed conditions as in the supply of human food. In the production of raw materials many are still following time-honored

but unscientific procedures; yet, in general, agriculture has become scientific and efficient through measurement. In the dairy industry alone the development of instruments for the measurement of butter fat in milk has stimulated the breeding of better cows and has made profitable dairying possible. The development of adequate measures for the detection of specific diseases in cattle, such as tuberculosis, has made it possible to produce milk which is free from danger to children.

Research in cereal husbandry, aided by measures, has helped to produce wheat which can be grown successfully in colder climates than any in which wheat had before been grown, and wheat which has more of the qualities necessary for bread-making.

Great changes have taken place during the past few decades in the preparation and preservation of foods for consumption in far-distant places or at some future time. Fresh vegetables are shipped from California to be consumed in Boston, and fresh fish is shipped from the Atlantic coast to be consumed in the Mississippi Valley.

More remarkable are the vast quantities of goods for canning and bottling. They range from all kinds of vegetables and fruits to meats, soups, jellies, dressings of all kinds, and delicacies from the far corners of the earth. The housewife buys her canned foods by "brand." The brand is the seal of the maker of the goods, a trade-mark which identifies the quality and flavor of the contents. The manufacturer must maintain this quality and flavor at all costs, and this can be done solely through the most painstaking and careful measurements not only of the amounts of flavoring (condiments, sugars, and so on) which go into the preparation of foods, but also through measurements of the strength, character, fineness, and like qualities of the flavorings themselves. The measurements taken in the production

of a certain brand of soup are the only guaranty of its continuance in popularity.

Measurement in education. As has been shown, progress in science and increase in knowledge in many fields, together with the numerous applications which have been made of this progress, have been due, at least in part, to the refinement of measures and to the increase in the effectiveness of measuring instruments. As there have been increases in the use of knowledge new and better measures have been found. With the introduction of new measures new instruments of measurement have been devised; with the standardization of new measures greater accuracy of observation has resulted and greater insight into possibilities; and with this greater observation and insight made possible, there has resulted new knowledge.

There is perhaps no field of modern thought where new knowledge is more needed than in education. Education — representing as it does the greatest single endeavor that modern man has been able to devise, in the pursuit of which there are the greatest number of leaders employed, and in which the greatest number of individuals are concerned, of all enterprises which have ever involved human energy — has been one of the last of man's activities to feel the urge of exact science. Speculation as to the proper means of education and question as to the means now in use are everywhere in evidence. Doubt concerning methods of teaching is as prevalent as doubt concerning the validity of materials of learning. The demands upon education and upon educators have increased beyond belief, and education, now commanding the attention of the whole world, must prove its truth or be replaced by truth.

The development of science through speculation and research, reënforced by measurements constantly increasing in accuracy and validity, has preserved and proved itself.

Pseudoscience (astrology, alchemy, magic, and phrenology), disproved by the same measures and by the same intellects which advanced science, has become buried in the ages that are past. Education that is truth will also prove itself, and education that is false will be abandoned, as have the other pseudosciences based on beliefs untenable in the light of advancing knowledge.

Such development can come only through measurement and through the refinement and use of measures to weigh the truths of education. Such measures are now available. The technics for using them and the justification for their existence form the theme of the following chapters.

Chapter summary. Progress in knowledge is partly dependent on the measures which man has at his disposal to determine his knowledge. For many years he has been handicapped by the crudity of his measures, which made full, complete, or accurate knowledge impossible. Such conditions resulted in superstition, guesswork, or quackery.

Progress in the development of human knowledge has been made through the increasing refinement of knowledge. In the past decade this acceleration has brought the old knowledge to new uses, and the newer knowledge to uses formerly undreamed of.

Modern methods of measuring and increased accuracy in measurement have brought extraordinary advances in the means of transportation and of the communication of ideas, and have made possible the mass production of foods, clothing, and the like on a scale never before possible.

What has taken place in these other fields of human knowledge and endeavor is now taking place in the field of education. Measurement and the means of measurement are of as much importance, and offer as much hope for the future, in this as in any other field of human interest.

CHAPTER II

THE BACKGROUND OF SCHOOL MEASUREMENT

Primitive measurement in education. In primitive life there were three main forms of living activities. The work of the men was largely concerned with providing for food supply and with warfare. The work of the women in primitive society was that of providing food, clothing, and shelter. In addition to these two activities there was a further and more highly specialized activity, that of the priesthood. Superstition following on ignorance of natural phenomena naturally led to the propitiation of unknown forces, and this in turn led to the necessity, or else created the opportunity, for the ritualistic leader, variously called priest, medicine man, exorcist, shaman, and the like.

The education in primitive life paralleled the living activities and was naturally centered in the family and in the home. The education of the boy consisted in definite preparation for the probable activities of hunting, fishing, and fighting. He was taught to make the objects he would need to use, as well as how to use them. He learned to fish by fishing and to hunt by hunting. He learned to make his implements for fishing and to construct his weapons for hunting. Not only were these things to be learned, but they were to be learned in definite and frequently highly prescribed ways — ways dictated by the traditions of the tribe and preserved by the priesthood.

This education, continued throughout the dependent years, culminated in the final ceremonies of admission, or

initiation, into the tribe. In a real sense this ceremony was the educational measure of the initiate. There were tests of endurance, fear, the ability to withstand hunger and thirst, bravery, accuracy, and knowledge of tribal customs. Sometimes practically the entire ceremony and nearly all the individual tests were included in one significant and powerful unit, as we find in the case of the African tuft-hunters; but more usually a ceremony lasting several days (during which the tests of endurance and bravery were given separately) constituted the measurements of that time.

The education of primitive peoples was simple. The objectives of that education were also simple, and this was so because life itself was simple. Custom and tradition supplied the motive for the living, and at the same time reduced the necessity for thinking. Thus living was a present occupation, bound up with the past in tradition and with the future in fatality. Therefore educational measurement was simple, objective, and direct.

Educational measurement in Greece. Both Athens and Sparta probably began their educational development in about the same primitive way, but their progress took two different courses: Sparta developed an educational system designed to achieve an immediate and practical end, whereas Athens sought to achieve a more remote and æsthetic end. Both civilizations developed side by side at approximately the same time and within a few miles of each other.

In Sparta the development of the individual was made subservient to the welfare of the state in all respects. To this end every effort was made to destroy the family as a social entity and to substitute for it the paternalism of the state. It is true that children spent their earliest years with their mothers, but even then they were hostages to the state, and the mothers were responsible to the state. After that time their lives and education were in the hands of the state

itself, and they lived at public expense in public barracks or dormitories. From then until the child became a man approaching thirty years of age, his education was a succession of militaristic trainings designed to inure him to hardships, to build up his physique, to develop military types of skills, and to exact from him the guaranty of implicit obedience and complete courage.

In Athens, particularly after the time of Pericles, emphasis was placed upon the development of the individual, but not necessarily a physical development for the protection of the state. Children were trained to run, jump, march, and the like for the purpose of developing them physically rather than for the purpose of developing their physiques for the use of the state. Combined with these physical arts was also the training in the æsthetic arts, as well as in other intellectual pursuits connected with oratory, grammatical study, and literature.[1] With these emphases on intellectual and æsthetic enjoyment of life, the education of Athens went beyond the strict utilitarianism of primitive peoples as expressed in an extreme by the Spartans.

In these two systems of living and of education are found the two contrasting elements of primitive education and cultural education. The Spartan education had been formalized and systematized until it had lost every semblance of individualism; the Athenian education had been freed until it became in the end almost completely individualistic.

Singularly enough the measurements of these two systems must have been startlingly alike. The test of the development of a Spartan boy was whether or not he could endure hunger, cold, and exertion; the test of the Athenian lay in determining whether he could run, jump, race, declaim, repeat literature, and debate. Both systems of

[1] Paul Monroe, *A Textbook in the History of Education*, pp. 102–117 ff. The Macmillan Company, New York, 1912.

education were based upon objective standards, and the measurement of these standards lay in the determination of how well they were attained. The difference between the two states lay in the fundamental difference in the objectives of the people, — service of the people to the state or service of the state to its people. These objectives governed the lives, education, and measurement throughout.

The measurements of chivalry. Under the influence of feudalism and the spread of Christianity a somewhat different educational development came about. This was the institution of chivalry. In a real sense this system of education combined many of the characteristics of both types of Greek education, although it added considerably to them.

Education in chivalry considered from the point of view of one destined to become a knight partook somewhat of the Spartan form of education, since the page was drilled in all types of exercises to fit him for fighting. Practice with sword and lance and practice in the many aspects of the care of armor, horses, and equipment while serving as page or esquire to some knight took up much of his time in preparation for knighthood. As with the Athenian, the æsthetic arts were not neglected, and music, literature, and deportment were also included. Beyond this, which in itself was not far different from former types of education, there was a definite effort to inculcate idealism, courtesy, respect, and care. These were expressed in the oath which the knight took on the occasion of his entrance into knighthood. This initiation had many of the characteristics of the initiation of youths into the tribes of primitive peoples. It was enhanced, however, with idealism and became a powerful influence over the lives of those who experienced it. As a final part of the ceremony the knight took an oath "to defend religion, to succor the distressed, to protect womanhood, and to be ever faithful to his fellow knights."

At this point measurement stopped. So long as the objectives of life were tangible, so long as they were objective, and so long as they could be determined or translated into definite activities or products, just so long were they measurable; after that the determination of the extent of the individual's education was a matter of guesswork, to be revealed only in his behavior in later years. Tests could be given for certain types of learning, such as skills of fighting or handling weapons, or for address, neatness, and the like, and, possibly, accurate judgments could be formed of such rather vague elements as honesty and courtesy in certain of its elements, but most of the objectives of education beyond these had to be accepted on trust.

Changes in the objectives of education. Education in the situations which have been considered was an activity carried on only with a selected few. In Sparta there were thirty or more times as many slaves and retainers who were not subject to the regimen of the education as there were citizens who were. A somewhat similar situation existed in Athens. During the Middle Ages by far the greatest proportion of the population was subservient to the remainder, who were nobles and rulers and from among whom came the small group of individuals who later became knights. It was not until the Crusades had decimated this small proportion of the population and had impoverished it at the expense of its retainers that the mass of the population became economically independent of their overlords. By lending money and goods to their overlords many of these people were able to buy their freedom, and in time whole "free" cities were established, where commerce and trade could be carried on without service to any feudal overlord save the king.

It was this situation that eventually resulted in the establishment of lower, or common, schools and that, continued and developed, has formed the foundation of the

curriculum which is now found in elementary schools. Education — the teaching to youth of the rudiments of common understanding and the tools by which intellectual progress may be made — spread slowly throughout the world, bound in rigidly at first by caste lines (even yet in certain localities), but becoming ever more universal, more standardized, and more effectual.

With the development of common schools there have come about many changes, both in the objectives of education and in the methods for attaining the objectives. The early schools were greatly restricted in their subject matter and in their scope. Their object was largely to teach children to read the Bible, to do simple numerical problems, and to write intelligibly. All these skill operations became easily organized, stilted, and formal, and the school itself displayed the same characteristics.

Formalism paved the way for a "faculty" psychology and the so-called "dogma of formal discipline." In the effort to justify the educational procedure and, at the same time, to develop an educational process that would be acceptable, all the forces at work were forces toward intensification of that formalism. The justification which was made was that education was a matter of training. This training was concerned with the mind of the individual. The mind of the individual, however, was made up of a number of separated compartments which acted as units. It was determined by experimentation and observation that there were certain brain centers of motor activities — of sight, hearing, smell, and taste — and that injuries to these portions of the brain resulted in changes or paralysis, which, in turn, affected sight, hearing, and the like.

From this idea of the localization of motor reactions there followed logically but mistakenly the idea that there were also centers for the higher forms of brain action or the results

of brain action. From this idea there came to be located certain brain centers of will, obedience, memory, and the like, and, in turn, certain forms of education became justified through their influence in developing these portions of the mind.

An early textbook in geography shows the extent to which formalism in the presentation of subject matter was carried in the schools. The entire text is in the form of questions with answers, but it is presented to the public as the desirable form of textbook construction.

Q. What are the situation and extent of New-England?

A. New-England is situated between 41 and 46 degrees of north latitude, and between 67 and 74 degrees of west longitude. It is five hundred and fifty miles long, and two hundred broad, containing eighty-seven thousand square miles.

Q. How is it bounded?

A. On the north by Canada; east by New-Brunswick and the Atlantic; south by the Atlantic and Long-Island Sound; and west by the state of New York.

Q. What are the civil divisions of New-England?

A. It is divided into the five states of New-Hampshire, Massachusetts, Rhode-Island, Connecticut, and Vermont; these states are subdivided into counties, and the counties into townships.[1]

In another section of this same geography the following is given:

Q. Give a concise description of London.

A. London is regularly built, contains about one million of inhabitants, and including Westminster and Southwark, is about eighteen miles in circumference. There are three hundred and five places devoted to religious worship in London, besides Methodist tabernacles, and exclusive of twenty-one out-parishes. There are also one hundred alms-houses, twenty hospitals and infirmaries, three colleges, ten public prisons, fifteen flesh markets, one for live cattle, two for herbs, and twenty-three others

[1] Nathaniel Dwight, *A Short but Comprehensive System of the Geography of the World*, pp. 138–139. Hudson and Goodwin, Hartford, 1802.

for corn, coals, hay, etc., fifteen inns of court, twenty-seven public squares, three bridges, forty-nine halls, eight public schools, one hundred and thirty-one charity-schools, two hundred and seven inns, four hundred and forty-seven taverns, five hundred and fifty-one coffee-houses, five thousand and seventy-five ale-houses, one thousand hackney-coaches, four hundred chairs, seven thousand streets, and one hundred and fifty thousand dwelling houses. There are no elegant royal palaces in England. Windsor Castle is the best.[1]

To justify a schematic system of teaching of this type the assumption which followed — that training in memory, for example, influenced all forms of learning where memory was involved or that training in attention influenced all forms of learning where attention was involved — easily led to the further assumption that it made little difference what the learning was, provided that it influenced the particular faculties decided upon as virtuous. Hence the next step in textbook construction was evidenced most clearly in the arithmetics and spellers.

In the arithmetics and spellers it made no difference what the problem was so long as it was a problem, nor whether the words to be spelled were useful, provided they could be spelled.

An early arithmetic shows this tendency in the following problems:

51. If the country which supplies the river Po with water be 380 miles long and 120 broad, and the whole land upon the surface of the earth be 62,700,000 square miles, and if the quantity of water discharged by the rivers into the sea be everywhere proportional to the extent of land by which the rivers are supplied; how many times greater than the Po will the whole amount of the rivers be?[2]

[1] Nathaniel Dwight, *A Short but Comprehensive System of the Geography of the World*, p. 33. Hudson and Goodwin, Hartford, 1802.

[2] Daniel Adams, *The Scholar's Arithmetic*, p. 193. Adams and Wilder, Leominster, Massachusetts, 1801.

9. Three jealous husbands with their wives, being ready to pass by night over a river, do find at water side a boat which can carry but two persons at once, and for want of a waterman they are necessitated to row themselves over the river at several times; The Question is, how those 6 persons shall pass by 2 and 2, so that none of the three wives may be found in the company of one or two men, unless her husband be present.[1]

The tendency in spelling is illustrated in Fig. 1, which shows a page from an early common-school speller.

The curriculum of the elementary school was narrow in the extreme, consisting to a large extent of the teaching of writing and reading. To the latter was added a simple form of arithmetic, grammar, spelling, and the like. Geography and history were later additions, and manual training, domestic science and domestic art, æsthetic appreciations, civics, music, and nature study are comparatively new. With this expansion of the curriculum there followed the necessity for the justification of each additional element, and this in turn increased the strain put upon the theory of faculties.

Although the theory of faculties, following upon a demonstrated fact of localized centers of physical attributes, was wrong, nevertheless some of the by-products of education which resulted from its practice were undoubtedly good. However, when the idea of faculties was carried forward to the next logical step, it became vicious. This step, which involved the notion that there was a correlation between the size of the supposed centers of will, determination, discipline, memory, and the like, and the degrees of will, determination, discipline, and memory exhibited by the individual, brought about the pseudoscience of phrenology and led to absurdities in education and related fields. In education, to develop memory almost any kind of material could be

[1] Daniel Adams, *The Scholar's Arithmetic*, p. 194. Adams and Wilder, Leominster, Massachusetts, 1801.

ACCENTED ON THE FOURTH AND FIFTH

ăr-²ĭs-tō-crăt′′ĭ-căl
cŏr-rō-sĭ-bĭl′ĭ-tȳ
dĭs-sĭm-ĭ-lăr′ĭ-tȳ
dĭ-vĭs-ĭ-bĭl′ĭ-tȳ
ĕc-clē-s ĭ-ăs′tĭ-căl
ēl-ĭ-gĭ-bĭl′ĭ-tȳ
ĕn-cȳ-clō-pē′dĭ-ă
ĕx-tĕm-pō-rā′nē-oŭs

ĭn-crĕd-ĭ-bĭl′ĭ-tȳ
ĭn-făl-lĭ-bĭl′ĭ-tȳ
ĭn-fē-rĭ-ŏr′ĭ-tȳ
ĭn-strū-mĕn-tăl′ĭ-tȳ
ĭr-rĕc-²ŏn-cī′lă-ble
măl-lē-ă-bĭl′ĭ-tȳ
mē-dĭ-ă-tō′rĭ-ăl
mē-tē-ō-rŏl′′ō-gȳ

spĭr-ĭt-ū-ăl′ĭ-tȳ
sū-pē-rĭ-ŏr′ĭ-tȳ
sŭs-cĕp-tĭ-bĭl′ĭ-tȳ
căm-ē-ră-ŏb-scū′ră
cir-cŭm-năv-ĭ-gā′tion
ĭn-tĕr-lĭn-ē-ā′tion
ĭp-ē-căc-ū-ā′nhă

ĭm-pĕn-ē-tră-bĭl′ĭ-tȳ
ĭn-cŏm-păt-ĭ-bĭl′ĭ-tȳ
ĭn-dĭ-vĭs-ĭ-bĭl′ĭ-tȳ

fă-mĭl-ĭ-ăr′ĭ-tȳ
gē-nē-ă-lŏg′ĭ-căl
hĕt-ĕr-ō-gē′nē-oŭs
hī-ē-rō-glўph′ĭ-căl
ĭm-pàr-t ĭ-ăl′ĭ-tȳ
ĭm-pĕt-ū-ŏs′ĭ-tȳ
ĭm-plā-că-bĭl′ĭ-tȳ
ĭn-cŏn-trō-vĕr′tĭ-ble

păr-ă-dĭ-sĭ′ă-căl
pē-cū-lĭ-ăr′ĭ-tȳ
plĕn-ĭ-pō-tĕn′tĭ-ă-rȳ
prăc-tĭ-că-bĭl′ĭ-tȳ
prē-dĕs-tĭ-nā′rĭ-ăn
pū-sĭl-lăn-ĭm′′ĭ-tȳ
rē-frăn-gĭ-bĭl′ĭ-tȳ
sō-c ĭ-ă-bĭl′ĭ-tȳ

pĕr-sŏn-ĭ-fĭ-cā′tion
răt- ĭ-ŏc-ĭ-nā′tion
rē-că-plt- ū-lā′tion
rĕc-²ŏn-cĭl-ĭ-ā′tion
sū-pĕr-ĕr-²ō-gā′tion
trăn-sŭb-stăn-t ĭ-ā′tion

ĭr-rĕf-²ră-gă-bĭl′ĭ-tȳ
mē-tē-ō-rō-lŏg′ĭ-căl
văl-ē-tū-dĭ-nā′rĭ-ăn

FIG. 1. Page from a school speller[1]

[1] Reproduced from page 97 of *A Pronouncing Spelling Book*, by J. A. Cummins. Published by James I. Cutler and Company, Bellows Falls, 1826.

memorized, to increase a skill almost any kind of skill could be developed, and to increase a "faculty" any portion of that faculty could be developed. When it was clearly demonstrated that there was no relation between the size of the brain case at any point and the amount of brain matter beneath it, phrenology as a science collapsed. When it was shown conclusively that the influence of skill in one field brought almost no similar influence in a related field, faculty psychology collapsed also.

The doctrine of specific trainings was then developed, and with it the entire field of behavioristic psychology. The first great advance here was an entirely new conception of the learning process. With the conception that learning was a specific reaction to a specific situation, and that general transfer of training occurred far more rarely than had been supposed, if it occurred at all, then the entire purpose of schools and schooling underwent a radical change.

The methods, characteristics, aims, objectives, and results of teaching became more closely examined, more carefully weighed, and more intelligently questioned. Where formerly, to justify a procedure, it had been necessary only to show that it was more or less related to some reasonably desirable objective, it now became necessary to justify each item of learning by itself and not by some fancied relationship. Thus there entered into the discussion of the place and value of the elements of the elementary-school curriculum one of the most dangerous but hopeful ideas. On the one hand, it had become theoretically possible so to standardize and formalize the elementary school as to mechanize it completely; on the other, it became possible to make broad inquiry into the really desirable objectives of school learning and to dignify them by making provision for them in teaching children. The technic of teaching has had to keep pace with these changes.

Both these movements have been in operation side by side for a generation, each acting as a check upon the dangers of the other. The one has brought into the elementary school a broad range of new subject matter and new emphasis which are characteristic of all new curricula in elementary schools. It has brought about new methods of teaching, has resulted in new attitudes toward children, has transformed the atmosphere of many schoolrooms, has provided for the expression of individual differences in the classroom, and has enabled many teachers to perform research, to experiment, and to grow. The other has resulted in a stabilizing of the content and placement of items in the curriculum. It has necessitated the careful scrutiny of new methods of teaching which have been proposed. It has made it necessary to justify new departures in terms of tangible assets. It has regulated the progress of children in terms of economic and social necessity, and it has prevented much erratic and unstable teaching from becoming widespread. Although it is possible to find in almost any school system extreme examples of each kind, — the informal, erratic, sentimental teacher and classroom on the one hand, and the formal, mechanized, and superorganized teacher and classroom on the other — nevertheless the two forces have in most instances harmonized one another, with the result that teaching and learning have become more free, less pedantic, and more effective.

The second great advance which can be directly attributed to the development of a rational behavioristic psychology is a new conception of the judgment of the efficiency of the learning process. Once the learning process was believed to be one of specific reactions to specific situations. When it was found that there were differences of all kinds, just as there are differences between individuals, a way was immediately opened toward the measurement of these learnings.

As has been shown, measurement of education up to this point had been largely a matter of final proof of ability, with much development remaining unmeasured. It was directly connected with the objectives of education, so far as they were developed; and what measurement there was, was in terms of the objective itself. The primitive boy made an arrow; the Spartan boy proved his endurance; the Athenian boy made a speech; the esquire demonstrated his skill with a lance; the boy of colonial days read, or figured, or spelled. So far as the objectives were valid, so far was the measurement good.

With the tremendous increase in objectives under the influence of behavioristic psychology, with the intense interest in locating and defining learnings, and with the added complexity surrounding mass learning because of the new knowledge of individual differences, the older form of measurement of final objectives became too crude and too unwieldy. School measurement had to be in terms of objectives, to be sure; but these objectives were too numerous and too complex for the few simple forms of measurement available, and the measures in use were too crude and unreliable to be of real service in terms of the new needs.

These new needs of education for refined measures and for effective means of measurement brought about the efforts (described in Chapter III) for supplying these measures and the means for using them, and also made possible, as will be shown later, the better definition of objectives, the clarification of individual differences, the delimitation of objectives, and the comparison of pupils, grades, and schools.

Chapter summary. Primitive education was in terms of the simple objectives and simple activities of primitive life, and the measurement of education was in terms of the objective standards of that life.

In Greece there developed simultaneously within a few miles of each other two forms of education. In Sparta, where the individual was subordinated to the state, the education took the form of definite trainings adapted to the preservation of the state. In Athens education was adapted to the welfare of the individual, and although the state was upheld as an ideal the education was primarily individualistic. In both cases, however, the measurements of the educational results were in terms of the objectives set up and were concrete expressions of the training received.

In the Middle Ages the education of the page and esquire to become a knight partook of the rigor of the training of Sparta and the more æsthetic education of Athens, with the addition of the inculcation of an idealism. In terms of concrete expressions of the objectives, the measurements of this education were on about the same level as those of the previous times.

With the changes that followed feudalism, resulting, as they did, in an independence of the people, and with the breaking down of caste lines, universal education became more possible, and common schools really became common. Because of its economic and social aspects this situation necessitated the justification of the education that was given, and that justification was finally found in the theory of faculties. Under this theory it was necessary only to show reasonable association of an educational activity with some educational objective in order to bring about its justification. This was in a sense the measurement of that education.

When it was conclusively shown that learning is a matter of specific reaction to a specific situation and that individuals differ inherently from one another, the place occupied by faculty psychology had to give way to a behavioristic form of thought. When this was established, two great changes resulted. The first of these was a new conception

of learning; the second was a new need for the measurement of the products of learning. The first of these resulted in a wide revision of the work of the school, whereas the second resulted in the development of measures, objectives of learning, and means of measurement which are considered in this book.

SELECTED BIBLIOGRAPHY

CUBBERLEY, E. P. The History of Education. Houghton Mifflin Company, Boston, 1920.

GRAVES, F. P. A History of Education. 3 vols. The Macmillan Company, New York, 1909–1913.

MONROE, P. A Textbook in the History of Education. The Macmillan Company, 1905.

PARKER, S. C. The History of Modern Elementary Education. Ginn and Company, Boston, 1912.

CHAPTER III

THE DEVELOPMENT OF SCHOOL TESTS

The traditional school measures. Tests or examinations have been the traditional forms of school measures. They have been utilized in both oral and written form and have appeared in many different ways. In the oral form they have been used as questions in the conduct of the recitation, and in the same form as the oral examination given to the candidate for the doctorate. In the written form they have been used in the traditional and well-known essay, or factual, type of examination, and in the same form as the thesis, or dissertation, presented as a result of research for the doctorate.

In most cases these tests have been devised by the teacher of the pupils. They have been given to the pupils by him, and the results have been judged by him. An exception to this procedure is found in the examinations set by the Regents of the State of New York and by the examinations set by the college-entrance boards and the individual higher institutions.

These tests have always served a useful purpose and still do so, but they are nevertheless subject to difficulties and defects that are hard to avoid. One of the gravest of these defects is the unreliability of the results gained from such tests. It might be thought that to a simple question set by a teacher a simple answer could be easily judged by that teacher. In some cases this is true, but in general it has been found by repeated experiments that it is not true. If, for example, the same examination paper is duplicated

many times and sent out to a number of different judges for correction, it is rarely that these judges will agree on their grading. This experiment has been tried many times and usually brings about much the same result.[1] A single paper in history was ranked by seventy teachers of history in high schools, where the passing grade was 75 per cent, with the result that the ratings ranged from 43 per cent to 92 per cent. In this case the judges had only the answer paper to influence them in making their gradings. What is true with respect to the difficulties of judgment of these history teachers is true in greater degree of the classroom teacher. The classroom teacher knows the pupil, the type of work which he has been performing, his social standing, and his general attitude toward his work, the other pupils, and his teachers. He is acquainted with the pupil's methods of work and believes that he knows his relative worth. No matter how diligent, sincere, or careful a teacher may be, it is almost impossible, if not wholly so, to prevent opinion on such matters from entering into the judgment of examination answers. Not only do different judges vary in their ideas as to the worth of such papers, but the same judge will vary his judgment on the same paper from one time to another. The teacher who will grade a series of examination papers, put them away for a period of time, and then regrade them without consulting or being influenced by his previous judgment will usually find very different results.

A second difficulty in the use of examinations of this type lies in the fact that the range of the tests is relatively narrow and therefore allows for inequalities in the answering as among different pupils, even though the judgments of the teacher do not vary.

[1] Daniel Starch and E. C. Elliot, "Reliability of Grading High-School Work in Mathematics," *School Review*, Vol. XXI (1913), pp. 254–259. Describes one of the earlier investigations in this field.

The relative restrictedness of the traditional school test may be shown graphically as follows: In an average test, say of four questions and with a group of four pupils, let it be granted that for a certain unit of subject matter each of the four retains a knowledge of 50 per cent of the materials involved, but a different 50 per cent in each case. On this basis, and also on a 100 percentile marking system, each of the four should receive the same mark, namely 50 per cent, since each knows the same amount of the materials. For the sake of argument it is also supposed that these four individuals can put on paper all they know, and cannot put down on paper what they do not know.

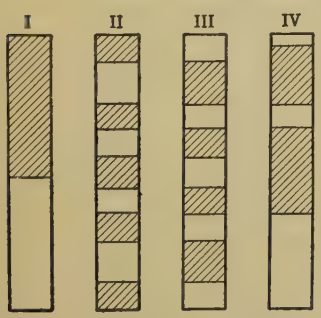


FIG. 2. Shaded portion shows distribution of knowledge of four pupils

If the four questions are distributed throughout the subject matter in the manner shown in Fig. 3, Question 1 involving for its correct answer a knowledge as indicated by the line shown, Question 2 as indicated by the next line below, and so on, it may be seen that pupil I is able to answer satisfactorily the first and second questions but is unable to answer the third and fourth. This gives him a mark of 50 per cent, which is his true mark.

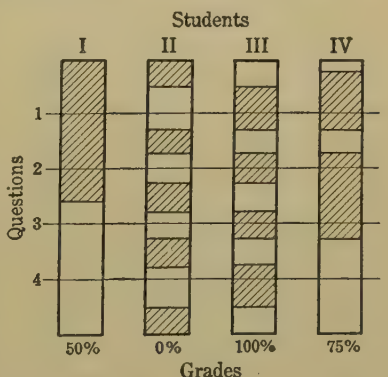


FIG. 3. Grading inequalities possible among four pupils having the same total amount of knowledge

On the other hand, with the same questions and with an equivalent amount of knowledge of the subject, pupil II is unable to answer any of the four questions and therefore receives a grade of zero, when

his actual mark should be 50 per cent. Pupil III is able to answer all the questions, although he knows no more, from an absolute standpoint, than pupils I and II. His grade is, nevertheless, 100 per cent. Pupil IV is able to answer three of the questions and is unable to answer the remaining one. Therefore his grade on this examination is 75 per cent.[1]

The real disadvantages inherent in these informal and traditional forms of examining lie within the tests themselves and in the means of scoring, and are not necessarily the fault of the teacher. The defect within the test is that it is a measuring instrument which varies greatly. If there were yards of widely different lengths, the meaning of a "yard" would be obscure and impossible to determine. So it is with these informal examinations. A stick of unknown length can be measured against a yardstick and its true length can be learned from that. Since the tests are variable in difficulty, and since there is no standard which does not vary with the difficultness of the test when it is rated on a percentile basis, every such test is in reality a different measuring instrument. The defect in the means of scoring lies in the crudity of the measures used. If the same element can be judged differently by two individuals, or even if the same individual can judge the same element differently on two occasions, then the measure used is crude and unreliable.

Beginnings in the refinement of measures. As has been previously shown, the real necessity for the development of more refined measures of educational products was brought about by the shift in emphasis from the belief in "faculties" to a belief in specific learnings. Hence, although there were some early attempts to find better measures, it was not until the new idea with respect to learning had become widespread that progress in the refinement of measures

[1] Charles Russell, Classroom Tests, pp. 14, 15. Ginn and Company, Boston, 1926.

could take place. Trabue [1] cites the attempt of the Reverend George Fisher of the Greenwich Hospital School in England to develop more adequate measures of educational achievement through a *Scale Book*, in which were placed samples of varying degrees of achievement of pupils. These samples of achievement, with sample examination questions and ratings of value, were to constitute the standards of achievement and the ratings of value in that school. Trabue also states that the report of this *Scale Book*, published by the Reverend Mr. Fisher, did not seem to have much effect on his contemporaries. It could hardly have been expected to do so in a day when education was being justified by the fancied transfer of its effects.

What was probably the first real effort to establish a better measure of educational products was initiated by Dr. J. M. Rice in an attempt to throw light on the solution of a controversy which was then being waged among educators. During the last decade of the nineteenth century, and particularly toward its close, there had been much agitation among schools for additions to the curriculum of the elementary school, and subject after subject had been added, with the result that less time could be devoted to the time-honored subjects. It was strongly urged by those opposed to this new order that the achievement of pupils in these important subjects was decreasing, and that more attention should therefore be given to them, largely through dropping nature study, physical training, manual training, and the like from consideration as proper subjects for elementary-school instruction.

In 1897 Dr. Rice constructed a spelling test, which he was able to give to many pupils in different cities where variable amounts of time were allotted to the teaching of

[1] M. R. Trabue, *Measuring Results in Education*, pp. 70–71. American Book Company, New York, 1924.

spelling. Attaining by the uniformity of his tests a uniformity of rating as among these cities, Dr. Rice was enabled to show that added increments of time in the study of spelling did not, as was being urged, bring proportionate increments in the achievement in spelling. As a matter of fact, he found that those schools which spent an average of fifteen minutes a day in the teaching and practice of spelling showed nearly as high achievement among their pupils as did those schools which required their pupils to spend forty minutes a day.

Dr. Rice's experiment of giving a uniform test to pupils in various localities was followed by other similar experiments, but in all of them there were defects in determining the real values. As in the case of the traditional examination, the main fault was due to the units of measurement and with the measuring instruments. The unit of measurement used by Dr. Rice could not be used except with the test that he had used and in the same way that he had used it. It was an improvement over existing measures, but its validity was confined to the schools in which it was used. All that Dr. Rice could determine from his units of measurement was that in City A so many words were correctly spelled, that in City B a different number were spelled, and that City B was exceeded by City A. In the second place there was no assurance that the "uniform" test which he gave was really uniform; that is, adaptable in equal proportion to the various localities in which he used it.

The establishment of reliable units of measurement. Dr. Rice supplied a new technic of measurement, but improved very little on the measures and measuring instruments then in use. It remained for Dr. E. L. Thorndike, then at Teachers College, Columbia University, to establish some reliable units of measurement, which he did in 1904, with the publication of his *Introduction to the Study of Mental and Social*

Measurement,[1] a pioneer work which is still standard in the field. The great contribution of this work of Dr. Thorndike was to make possible the uniform treatment of such results as had been found by Dr. Rice and others, to the end that uniform judgments could result. This establishment of a body of technic for the statistical interpretation of data in the educational field was the first step in the establishment of reliable units of measurement and has been the basis for all further refinement of those measures. So long as methods of interpretation of data varied it was impossible to find common ground for the comparison of data, but with the introduction of these technics it merely remained for educational investigators to establish reliable instruments.

The development of measuring instruments. When Dr. Rice made his investigation of spelling he constructed a test of spelling as his measuring instrument. As was pointed out, this test may or may not have been equally applicable in the various situations in which it was used, and therefore had little validity beyond those points. With the introduction by Dr. Thorndike of standard statistical means of defining units of measurement it became possible to make improvements in these tests; but it was long afterward that the possibility of a scale became realized in practice, and even now the difference between a test and a scale is frequently not clear. This is the reason that up to this point the terms "measuring instruments" and "units of measurement" have been used to indicate two different necessities in measurement. *A test is a measuring instrument* analogous to a yardstick or a foot rule; *a scale is made up of units of measurement* which are analogous to the divisions on a yardstick or a foot rule. Since the two are frequently found together they are thereby confused.

[1] E. L. Thorndike, *Introduction to the Study of Mental and Social Measurement.* Teachers College, New York, 1904.

Dr. C. W. Stone was probably the first to apply Dr. Thorndike's suggestions in the making of a test, and in 1908 he published the Stone Reasoning Test in Arithmetic. Dr. Stone constructed (for sixth grade only) a test containing a number of problems in arithmetic which required reasoning in their solution. He was interested mainly in determining the difficulties of sixth-grade pupils in reasoning; and through giving the test to many pupils he was able, by means of Dr. Thorndike's statistical methods, to determine the relative degree of difficulty of the questions which he used. When the test was published, it was intended to reveal the ability of a sixth-grade pupil to reason in problems in arithmetic, and evidently no thought was given to its use as a scale. It represented, however, a great advance in test construction, and was the first attempt to weight the values or difficulties of the elements in a test.

The next advance was made by Dr. Thorndike himself in 1910, when he published the Thorndike Handwriting Scale. This was the first real scale for the measurement of the products of education. It consisted of a large number of samples of the handwriting of children arranged in the order of excellence as estimated by competent judges and with each sample allotted a numerical value. By means of this scale it was made possible for a teacher, or even a pupil, to compare a sample of handwriting with the various samples given on the Thorndike Scale and to judge which sample was most nearly like the work of the pupil, thereby giving the work of the pupil a rating of excellence in his own work. In presenting his scale Dr. Thorndike pointed out that the writing of two different individuals could be compared through this scale without comparing them directly (for example, a child in one community with a child in another), or that successive efforts of a single individual could be compared to measure improvement.

This development of a scale was followed by other scales brought out under the leadership of Dr. Thorndike, notably the Hillegas Scale for the Measurement of English Composition by Young People in 1912 and, later, the Buckingham Scale of Spelling Ability. Both these scales reproduced, in different fields of subject matter, the essentials of the Thorndike Handwriting Scale, in that they arranged the samples of work indicating increasing degrees of accomplishment or difficulty in such a way that relative values could be assigned to the various merits of accomplishment.

The development of these tests and scales paved the way for the next advance in school measurement, since they provided materials necessary for what was to follow. On the one hand they provided a technic of scale construction, and on the other they refined materially the methods for the construction and use of the tests themselves, particularly the methods of scoring and judging the quality of the results of the effort of pupils.

The establishment of standards or norms. After the publication of Dr. Stone's work in arithmetic reasoning, tests for the measurement of arithmetic accomplishment of school children were made by Dr. S. A. Courtis of Detroit, to whom credit must be given for the next great advance in the development of school measurements. Dr. Courtis made tests which were designed to measure the rate at which children were able to do the four fundamental operations in arithmetic. He administered these tests to several thousand school children in different states, and in 1911 published the results which he had obtained in an article entitled "Standard Scores in Arithmetic."[1] It is significant that Dr. Courtis used the word "standard" and that he gave the same tests in all the different grades of the elementary school from the

[1] S. A. Courtis, "Standard Scores in Arithmetic," *The Elementary School Teacher* (November, 1911), Vol. XII, pp. 127–137.

third grade through the eighth. The establishment of standards of accomplishment in the same field, indicating degrees of educational advancement in that field on the part of school children, using the same measure for all the pupils, and thereby making a standard of accomplishment for the various grades, was a big contribution. Later investigations needed only to establish the validity of the standards and to locate them for other tests and scales. In this way a new element had entered that not only determined the actual accomplishment of pupils but also provided a standard by which to compare and judge the value of the accomplishment. It also foreshadowed what was perhaps the most important contribution for the new scales (which were at this time independent of the tests or test materials that were used); namely the age scale, the grade scale, and the T scale.

Shortly after the publication of Dr. Courtis's article on standards he was invited to apply his tests in a general survey of school conditions which was being undertaken in New York City. In this connection he applied the technic of testing and of norms, or standards, which has since become one of the fundamental factors in the survey of an educational institution.

The development of the measurement of intelligence. It is a strange coincidence that at the very time that these developments were taking place in the field of educational measurement, a similar development, independent of them, was occurring in a different but later related field — that of mental measurement.

For many years efforts had been made to find some measure of intellect, or intellectual ability. Various measures were proposed from time to time, and there was a certain body of incomplete knowledge at hand in 1904, when a French psychologist, Dr. Binet, and a French physician, Dr. Simon, together undertook to find some measure for the

location of subnormal children for special educational classes. As a result of their efforts (in which they combined much of the knowledge and investigation which had preceded them and to which they added ingenious insight) they were able in 1905 to publish a tentative scale for the measurement of intelligence, primarily for the purpose of locating subnormal children. In 1908 they published a revised, extended, and improved scale known as the Binet-Simon Scale for the Measurement of Intelligence. In this scale were included a large number of tests of various types which were thought to show variations in mental ability.

In 1910 Dr. H. H. Goddard made a translation of this scale for use in America and advocated it, as had Binet and Simon, for the location of really subnormal children in need of special forms of education. This scale was found to be poorly adapted for American children, inasmuch as it was too easy in the lower ranges and too difficult in the upper ones. After much investigation the test was rearranged, added to, and published by Dr. Terman as the *Stanford Revision of the Binet-Simon Scale.*

This revision was widely spread throughout the United States during the years that followed, but it had a serious defect which confined its use to a relatively few outstanding cases: it was an individual test, and the examiner had to spend his time with a single child. Since the application of the scale necessitated a skilled examiner, and since the time of skilled examiners was limited, it could be applied only to a small proportion of the school population.

With the entrance of the United States into the World War and the consequent mobilization of thought and action toward efficient and rapid participation, the use of intelligence tests to determine fitness for various assignments of soldiers became necessary. As a result a number of eminent psychologists, mindful of the value as well as of the limita-

tions of the individual intelligence test, combined their efforts to produce a test which could be given to the large numbers of soldiers with a limited force of examining leaders. Out of this there resulted the Army Intelligence Tests: the Army Alpha, for those who could read and write, and the Army Beta, designed for illiterates. These tests were designed to measure much the same abilities as were measured by the Binet-Simon Tests, but they were adapted and arranged in such a way as to make it possible to administer the tests to large groups of individuals at one time; hence the name "group tests."

The success of this technic of group-testing inspired others to adapt it to the use of schools, and there resulted two sets of intelligence tests for school children: the Otis Tests of Intelligence and the Haggerty Tests of Intelligence. Shortly afterward there appeared a modification of the Army Alpha Test, called the National Intelligence Test, which was intended for use with children in those grades where there was reading ability. Other tests adapted for the same and other purposes followed, but all aimed at the measurement of intelligence. Some of these tests were for use with nonreading pupils in primary grades, others were for high-school and college students not measured sufficiently by the group tests first developed, and still others were for measuring definite forms of mental ability rather than general forms. With the development of these tests and with their consequent rapid spread in use it soon became possible to measure and compare achievement and ability in the same individual, thereby completing another important step in the development of school measurement.

The development of classroom tests and testing. The two defects in classroom tests — the restricted range of the tests as they were made and the nonobjectiveness of the judgments of the answers — had been largely eliminated through

the use of standard tests. However, standard tests were not available in sufficient variety and breadth to satisfy the needs of the classroom teacher, with the result that technics for achieving greater objectivity in scoring and greater range in tests made by the teacher were needed to add to his measurement technic within the classroom. On the basis of technics already established in standard testing, efforts were made to improve the two main defects in teachers' tests or informal examinations.

To provide a greater range of material for the tests, the first adaptation was the inclusion in a teachers' test of a large number of elements each of which required a smaller amount of time for answering than did the traditional examination question. One of the earliest attempts resulted in what is now known as the true-false test, in which the pupil is confronted with a number of statements and is asked to denote those that are true and those that are not true. Other types of tests followed which were designed to measure other abilities, such as discrimination of relevant from irrelevant materials, the placement of items in chronological or logical sequence, the matching of items related to one another, and the like. A body of technic for classroom testing developed which made measurement of classroom products by the teacher a more useful and accurate procedure.

The increase in objectivity was largely accomplished through devising test elements which could be answered by a few words at most and which could be scored without variation by the same individual. Some of these tests were more objective than others, but in all types the necessity for judgments by the teacher were reduced as far as possible.

Classroom tests, devised first by psychologists for college students and collegiate types of subject matter, were adapted for use in high schools and elementary schools and were followed by means of utilizing the results of the tests for

greater interpretation of pupils' work. Technics were devised for determining the relative difficulty of tests, for equating the difficulty of test parts when tests were used in batteries, and for analyzing the results of tests to diagnose the difficulties of pupils or for the improvement of teaching. On the basis of the use of achievement ratios in standard testing, it became possible to devise a similar effort score for use in classroom testing and also to devise means whereby a score on one test could be made comparable with a score on another.

All these technics made possible the use of classroom tests for the improvement of grading and marking and increased their value as a teaching device. The labor of testing pupils has been shifted from the interpretation of answers to questions to the making of better questions and the interpretation of results, an improvement which is in itself of great and lasting value.

The spread of testing. With the developments in the construction of reliable measures, or tests, in standard testing, in the development of workable and understandable units of measurement, or scales, and in the availability of accurate group measures of ability, testing became an almost indispensable tool in the hands of the teacher and administrator. As has been shown, the first developments came about in the field of elementary-school subject matter. After the early tests were made, others followed in rapid succession. A revision of the Hillegas Composition Scale (showing a larger number of points on the scale), called the Nassau County Revision of the Hillegas Composition Scale, was made. The Woody Arithmetic Tests were developed, and the Thorn Reading Tests, the Burgess Tests of Silent Reading, the Woody-McCall Arithmetic Tests, the Courtis Standard Practice Tests in Arithmetic, the Thorndike-McCall Reading Scale, the Morrison-McCall Spelling Scale,

and many others in the field of elementary-school instruction. Wherever there was a more or less universalized body of skill or knowledge the standard test was applied, and in some instances there appeared batteries of tests capable of measuring wide sections of elementary-school work.

From these beginnings testing spread both into primary and into secondary schools. Tests were constructed for the primary grades and for nonreading children. They were introduced into the high schools for use in English, Latin, grammar, science, and the like, and from there have gone on into the college field.

The practicability of the testing technics and the concreteness of the results have appealed to industrial and vocational leaders, and tests have been developed to reveal mechanical and vocational aptitudes, manual as well as mental alertness, and the adaptability of individuals to various forms of work.

At the same time that tests have been spreading throughout the educational system, they have become increasingly more useful, and the technics for describing their results have become likewise more reliable. They have given the teacher a means of locating the difficulties of his pupils in the learning process, they have enabled teachers, supervisors, and administrators to define more accurately and completely their expectations of the development of pupils, and they have supplied an accurate and useful knowledge of pupils which was never before available.

Tests are now used not only to show what pupils have done but also to show what they might do. They are used as standards of attainment of pupils and as goals for effort. By combining the use of different tests a general survey is possible for determining the status of pupils for grading, promotion, and classification. Through the analysis of test results and through the proper selection of tests for subse-

quent use, the difficulties of pupils in learning or the deficiencies of teachers in teaching can be determined and acted upon. In short, although they cannot replace teaching or act as a substitute for it, tests have become one of the most valuable tools available for the use of a teacher.

Principles of school measurement. The development of school measurement, as already indicated, has not been due so much to the zeal of educational investigators or inventors delving into an unknown and undiscovered region of education as to the increasing skill in the interpretation of the principles upon which school measurement rests. Dr. McCall has stated some of these principles in the form of theses in his *How to Measure in Education.* However, since they are fundamental to an understanding of not only the present of school measurement but its future as well, it is wise here to treat some of them as axioms in educational measurement. At first glance they appear as truisims, but a deeper inquiry will reveal their implications.

1. *Whatever exists at all exists in some amount.* There can be no quarrel with the truth of this principle so long as concrete objects are considered. A horse, since it is a thing, exists in some amount. It stands so many hands high, it has so much weight and so much mass, and the like. It exists, and it therefore exists in a definite amount. A like proof can be given for a tree, an orange, a building, or a person.

When something other than a concrete object is considered, a different form of proof must enter into consideration. For example, to what extent are there amounts of knowledge of history, amounts of knowledge of geography, and amounts of knowledge in general? Here it can be said that if there is a knowledge of history it exists in some amount. Of the elements of the history of the Greeks there is much present knowledge, and the amount of that knowledge increases

with each new manuscript that is found and with each new excavation that is made. Of the elements of the history of the Aztecs and of the civilization which preceded theirs, knowledge is far more fragmentary, but it must be clear that such knowledge as there is exists in some definite amount.

Each individual possesses certain skills (for example, in walking, running, reading, drawing, handling a fork or a knife, and so on), and these skills are present in each individual in some definite amount, although the amount may vary in different individuals.

On other levels, although one may be willing to admit of the existence, it is more difficult to prove that the elements exist in some amount. The attraction between individuals, if it exists at all, exists in some amount. Such elements as honesty, reliability, tact, and responsibility, if they exist at all in an individual, exist in some definite amount. Thus there may be degrees of attractiveness or repulsiveness, of honesty or dishonesty, of reliability or untrustworthiness, of tact or carelessness, of responsibility or thoughtlessness. Whatever exists at all — be it concrete or abstract, quantitative or qualitative, apparent or obscure — must exist in some amount. If it has no reality, that is, if it does not exist, it has no amount.

2. *Whatever exists in amount can be measured.* This principle has been greatly misunderstood. There are many things which exist in some amount which have not yet been measured. There are also, let it be hoped, many things which exist in some amount whose presence has not yet been suspected, to say nothing of being located and measured. It is possible too that there are many things which it is undesirable to measure. All that this principle states is that whatever exists in some amount can be measured.

The chemists who, with great ingenuity and skill, have located and measured the atomic weight of rare earth metals

which previously were unknown have been working directly with this principle. They have reasoned that the metals existed. They did not know what they were and possibly they did not know where to find them, but something which did not exist within their knowledge was revealed in a spectrum. That proved the existence. Then followed the search for the material that would indicate the presence of the same something which was revealed in the spectrum. With that located, there followed the experiment and trial and measurement which would reveal the material itself. Pasteur made his great advances in his work by dealing directly with this principle. He recognized the fact that a disease of silkworms existed. He assumed that the cause of the disease was real and that it existed in some amount. His efforts were then directed toward the location, or measurement, of that amount, and with this located he sought a means to remove the cause he had found.

This principle has been followed strictly in the establishment of both tests and scales in education. Assuming that arithmetic ability existed in some amount, Dr. Courtis made his efforts to measure it. Assuming that ability in handwriting existed in some amount, Dr. Thorndike tried to devise a means of measuring it. Assuming that ability to write English composition existed in some amount, Dr. Hillegas attempted to measure it. Every successful test or scale in the educational field has been based upon the principle that if a thing existed in some amount it was measurable, provided the proper measure could be devised for the purpose.

3. *Efficiency in teaching depends upon desirable changes in children.* An analysis of the meaning of this principle is an analysis of educational effort. Whereas at one time it might well have been thought that the duty of a teacher was to give something to pupils, much as a textbook might present

to pupils an array of facts, the new knowledge of the learning process indicates clearly that this portion of the teacher's work is merely supplementary to all the other materials of teaching. It is now known that the learning process is one by which neurone changes of some kind occur in the child, and that these neurone changes constitute in the aggregate what is known as learning. It is also known that some of these changes are desirable in terms of accepted educational objectives and that others are undesirable. It therefore follows that the work of the teacher, which cannot be successfully done by any other agency, is not only to induce these changes in pupils but also to select desirable changes. The extent to which he is able to effect desirable changes, then, other things being equal, is the measure of his efficiency.

4. *The evaluation of learning depends upon the ability to make changes, as well as upon the changes made.* It is known that individuals vary as much in mental ways as they do in physical ways, and hence that the traditional standards of achievement which resulted in the idea that children should accomplish a "grade a year" and in the idea of "failure" in school work are wrong. That which a pupil should accomplish in the learning process should be determined not by some arbitrary standard set up for him by some outside agency, but rather by the definite standard which is set up by his own ability. Since it is the ability of a child which really determines the extent to which he is capable of change (in most cases an ability in terms of rate rather than kind), it should be clear that the evaluation of that learning should be in terms of that ability. The child who moves slowly in making desirable changes and does so because of an ability which makes slowness necessary, is doing satisfactory work, despite the standard that may be set up for any other individual. The teacher who effects those changes is also efficient. The child who moves slowly and whose ability would

indicate that his true rate should be faster is doing unsatisfactory work for him, and to that extent the teacher who fails to accelerate the learning is inefficient.

To carry this principle a step further necessitates the idea that comparison of classes or teachers or school systems in terms of achievement should also take into account the ability of the pupils. In a certain school system where school savings are much encouraged, the local newspaper publishes each week the total savings made by the pupils of the various schools. Each week one school heads the list with the total of savings greater than any other school, and each week implied praise is given to that school. In point of fact another school, which never has received praise for the efforts of its pupils to save, is probably deserving of more because, in terms of the ability of its pupils to save, it is saving more than the first school. The money saved by the pupils in the first school, which is situated in a wealthy section of the city, is largely given to them by parents who are financially well-to-do. In the other school the money saved is the result of cutting lawns, delivering papers, and other earnings on the part of children of weekly wage-earners. In terms of saving according to ability to save, the second school is far in the lead, and justice requires that the true evaluation of such saving should take this fact into account.

5. *The value of a desirable objective in education is determined by its tangibility.* Unless an objective in education can be defined in such a way that the learning changes necessary to achieve it may be found, and that when the changes have been made the objective can be recognized, it is of little value. An abstract honesty may be thought of as a desirable goal in education, but honesty is not abstract: it is made up of great numbers of small concrete reactions to as many concrete situations. Some of the reactions that may be called honest are not only specific but narrow in

character. The situation "apples in a store" with the reaction "pay for one if I want it" or "refrain from taking one even if I were in no danger of being detected" is but one example of a specific situation included in the idea of honesty. There are other situations that bring reactions which are more general in character and which have more widespread significance. The situation "something which does not belong to me" and the reaction "I must not take it without either (1) paying its value for it or (2) asking the owner to give it to me" has this general significance. To the extent that these desirable reactions included in the idea of honesty can be identified and made tangible they are of value as objectives; to the extent that they cannot be identified and made tangible they are worthless.

Not so long ago almost every objective in school work was to a large extent intangible. What constituted objectives in reading, spelling, arithmetic, and the like were almost unknown. With the refinement in measures, and with the definition of objectives in tangible form, great advances in all these fields of teaching have been made. Objectives in spelling can now be defined as so many words taken from such and such a list in the Thorndike Word Book, to be mastered in such and such a grade by pupils of normal ability; objectives in arithmetic can now be defined as the ability to solve correctly in so many minutes so many problems with such and such characteristics. This is tangibility, and there can be no question when the accomplishment of the objective has been reached.

The objection may be raised that there are objectives in education which are intangible and yet which are valuable. If so, what are they? Character, honesty, self-realization, initiative, freedom, responsibility, and a host of other so-called objectives in education have been advanced to justify types of learning that are going on in our schools. No one

would be rash enough to dispute the fact that these are desirable attributes for every child to have, and it may be that these are, and of right ought to be, the primary objectives of education. The extent to which they are valuable objectives, however, is limited by the extent to which, as objectives, they are really tangible. Character will be no less valuable if what it is and when it is reached is known. Arithmetic was no less valuable when its elements were defined and when each element was striven for conscientiously. Initiative and responsibility are in the end made up of many elements, each of which must be defined and made tangible, so that they may be specifically striven for and that they may be recognized when reached. Tangibility is not a means of mechanizing ideals or of deflating worthy aspirations. A tangible ideal is more likely to be reached than one which is so vague as to be unrecognizable.

6. *The worth of the methods or materials of instruction is unknown until their effect is measured.* Learning has ever been surrounded with devices and helps. From the introduction of the hornbook as a device in learning to the present day, when maps, books, pictures, diagrams, texts, outlines, and materials of every kind and nature are in use, it has been assumed that these materials have helped in the learning process. Do they help? Is the counting of pegs or sticks a help to the child who is learning to add integers? Or may it be a real hindrance? Are the things that help an adult in understanding the same things as those that help a child? Do the pictures in the geography textbook really add to a child's understanding of the study? None of these questions can be answered intelligently until the effect that they have on learning is measured. Every person may have an opinion, and most teachers do, but that opinion is based on a measurement which is crude and unreliable. Does the picture of an elephant aid a child in knowing what an elephant

really is? One child at least was not so helped, and for years gave the picture one name and the animal in the zoo or the circus another.

Does the map help in understanding territorial divisions? It seems certain that it does with the adult; but it is also true that there are many children who would be surprised to learn, like Tom Sawyer, that Connecticut was really green and brown rather than pink.

Is a movable desk, a fixed desk, or a table the best adjunct of a classroom? There are teachers who believe in one type, and there are those who believe in another. Only a valid measurement can decide which contributes most to learning.

One of the most-discussed problems in modern education is that of teaching-method. Shall it be by topics or by problems? in socialized form or in "project" form? Is formal teaching doomed? Is the Winnetka plan or the Dalton plan better than that in other school systems? Such questions as these can only be answered definitely by measurement of their effects. If they result in reaching the desired objectives and if they reach them more easily, more quickly, or more effectively, they are better. No dissenting opinion can controvert facts when facts are available. Dr. Rice measured the result of using different lengths of time in the teaching of spelling and found his results contrary to the opinion of teachers; but he had facts which the teachers did not have, and his judgment based on those facts was worth more than the collective opinion of all his opponents.

These principles do not constitute all those on which modern measurement is progressing; but they do constitute the more fundamental axioms which are responsible for the development of measurement as it is now developing, and it is upon such principles that the values of measurement are based.

Chapter summary. The traditional school measure is the familiar school examination, or test. In the oral form it ranges from the question in the recitation to the oral examination for the doctorate. In the written form it ranges from the written-essay type of examination to the dissertation for the doctorate. The major difficulties in the use of these traditional forms of tests lie, on the one hand, with the nonobjectivity in giving and scoring the tests, and on the other, in the fact that the range of such tests is relatively limited. These difficulties are inherent in the form of the tests themselves and are not necessarily due to the fault of the teacher.

Early attempts to find more reliable measures of school products were limited in their influence, until the Rice test in spelling established the possibility of valid comparative ratings of the products of pupils. Dr. Thorndike followed with the establishment of uniform units of measurement through the publication of his volume on mental and social measurements in 1904. In 1908 Dr. Stone produced his reasoning test in arithmetic, the first real standard test, and this was followed by Dr. Thorndike's scale of handwriting, which constituted the first real scale. Other tests, built on the knowledge and skill gained in these early tests, followed in rapid succession. The first use of standards, or norms, was made by Dr. Courtis, who in 1911 reported on the use of his arithmetic tests with many thousand children in several different cities.

Independent of this development, but merging with it in its later stages, was the development of the intelligence test. This was originated by Dr. Binet and Dr. Simon in France, who in 1905 published the preliminary form of their now famous individual intelligence test, a revision of which was published in 1908. Working on this model Dr. Terman, through the stimulation of Dr. Goddard, revised the Binet

tests for use with American school children and published it as the *Stanford Revision of the Binet-Simon Scale.*

During the World War a group of psychologists produced two group tests for the measurement of the intelligence of soldiers: one, called the Army Alpha, was designed for those who could read; the other, the Army Beta, was designed for illiterates. With the Army Alpha as a model, group tests were devised for the measurement of the intelligence of pupils in schools, and these have been widely used since that time.

With the construction of reliable tests, the development of scales by which to measure them, and the development of tests of ability, the spread of testing through the elementary school was rapid. Tests were made for high schools and colleges and were adapted for use in vocational guidance and other fields.

The wide use of tests today, the increasing number of useful tests, and their adaptability for use in supervision, diagnosis, remedial teaching, and the like are largely due to certain principles which have guided the entire development. These can be stated as axioms of testing.

The first of these axioms is that whatever exists at all exists in some amount. When this is applied to concrete objects its implication is clear; but the truth of the axiom is just as great when it is applied to various degrees of knowledge or to various degrees of skill or even to various degrees of ideals and appreciations.

The second axiom is that whatever exists in amount can be measured. This does not imply that whatever exists has been measured or even ever will be measured. It merely implies that for things which exist in different amounts, given sufficient ingenuity, a measure is possible.

The third axiom is that efficiency in teaching depends upon desirable changes in children. This is just another way of

stating the principles underlying the learning process; which imply that it is the teacher's responsibility so to direct situations as to produce learning in the pupils, not merely to give something to them.

The fourth axiom is that the evaluation of learning depends upon the ability to make changes, as well as upon the changes made. This axiom implies that ability as well as accomplishment must be taken into account in judging the degree to which learning is satisfactory from the point of view of the school. It means that the traditional units of learning-time must be disregarded for pupils, even though the units themselves may be justly kept as a basis for the learning of all pupils.

The fifth axiom is that the value of a desirable objective in education is determined by its tangibility. With some educators this has been a moot question for many years; they hold that some desirable objectives ought to remain intangible to prevent the mechanization of learning. Since tests need not mechanize learning, since the value of objectives in those fields where the objectives have been made tangible has been greatly increased, and since progress is being made in education largely through the adherence to this statement, it seems clear that objectives gain through becoming tangible rather than lose.

The final axiom is that the worth of the methods or materials of instruction is unknown until their effect is measured. Opinion has largely determined what materials shall be used in elementary-school instruction and when those materials shall be used. It has also been used to further or retard proposals and attempts in educational methods. This axiom merely states that the worth of these proposals and attempts cannot be determined through opinion, but rather through the definite measurement of their effect upon pupils.

SELECTED BIBLIOGRAPHY

GREENE, H. A., and JORGENSEN, A. N. The Use and Interpretation of Educational Tests, chaps. i–ii. Longmans, Green & Co., New York, 1929.

MCCALL, W. A. How to Measure in Education, pp. 14–16. The Macmillan Company, New York, 1923.

PINTNER, R. Intelligence Testing, chaps. i–iii. Henry Holt and Company, New York, 1923.

RUSSELL, C. Classroom Tests, chap. ii. Ginn and Company, Boston, 1926.

TERMAN, L. M. The Measurement of Intelligence, chaps. i–ii. Houghton Mifflin Company, 1916.

THORNDIKE, E. L. The Measurement of Intelligence, chap. i. Teachers College Bureau of Publications, New York, 1927.

TRABUE, M. R. Measuring Results in Education, chap. iii. American Book Company, New York, 1924.

PART II. TESTS

CHAPTER IV

TYPES OF TESTS

Standard and classroom tests. School measurements, as has been pointed out by Gregory [1] and McCall,[2] include more than the tests and scales which have been considered in Chapter III, but the school measures which can be considered of prime importance in their relation to the learning process and the progress of pupils through school are the tests and scales. These measures, particularly the tests, vary greatly in their characteristics, and it is well-nigh impossible to find mutually exclusive forms of classification to describe them.

A preliminary classification is not difficult, for between standard and classroom tests there are easily distinguishable and mutually exclusive differences.

One characterization of standard tests is their objectivity. Objectivity is the quality that measures the extent to which the elements of the test can be universally reacted to in the same way by the teacher or the scorer. A completely objective element would be one to which all persons would react in the same way in rendering a judgment and to which the same person would react in exactly the same way at two different times. All standard tests are not completely objective in this sense, for there are certain tests where there

[1] C. A. Gregory, *Fundamentals of Educational Measurement*, pp. 232–253. D. Appleton and Company, New York, 1922.

[2] W. A. McCall, *How to Measure in Education*, pp. 9–10. The Macmillan Company, New York, 1922.

may be some doubt on the part of the scorer as to the proper rating to allow for a given answer. However, these cases are rarer now than formerly, when the technic of constructing tests was not so well understood.

A second characteristic of standard tests is their comprehensiveness within the field measured. It is constantly being emphasized by the makers of standard tests that justice in comparisons of two or more communities on the basis of these tests depends upon the extent to which they are comprehensive, — the extent to which they apply equally to the needs of the different communities which they are designed to serve. A lack of a given content in a test would react against a community which emphasized that group of elements and might act positively for that group which neglected it. Standard tests are so constructed, through the selection of the elements which they contain, that they may be as comprehensive within the field to be measured as it is possible to make them.

A third characteristic of standard tests is the universality of their content. A field of teaching which is localized or provincial in character cannot be placed successfully in a standard test. Theoretically such a test could be standardized for a given restricted community where the localized subject matter was used, but actually the success of standard tests lies in their almost complete use of universalized subject matter. A test in arithmetic, for example, should utilize the subject matter in arithmetic of admittedly universal teaching. Arithmetic is a unit of subject matter that is widely distributed in much the same way and is therefore peculiarly adapted for use in standard tests. Other subject matter in the schools — geography and history, for example — have far less universality of content and are therefore to that extent restricted in the values of standard tests which can be utilized for them.

A fourth characteristic of standard tests is the fact that they have norms, or standards, of achievement. These norms are derived from the actual results of giving the tests to many pupils. The results thus obtained make possible an estimate of the achievement which might be expected of other and untested pupils.

As was shown in Chapter III, classroom tests in their present form arose through the need of measures more reliable than the former traditional essay types of examination, and are constructed and used somewhat in terms of the technics used in standard testing.

One characteristic of classroom tests is that they are usually made by the classroom teacher. There seems to be a slight tendency on the part of authors and publishers to place on the market printed forms of tests without norms, resembling classroom tests. It is possible that in many cases they serve a useful purpose, but it is also probable that those which are useful must eventually become adjusted through the gathering of norms and the careful revision of the elements they contain so as to become standard tests. The printed form of the classroom test, designed for wide distribution, if successful in that distribution, is in a preliminary form of standardization.

A second characteristic of classroom tests is that although they employ to a large extent the technics that make for objectivity in standard tests and are therefore much more objective than the traditional form of essay-type examination, they are far less objective than standard tests, since they depend to some extent on the teacher's judgment for their scoring.

A third characteristic of classroom tests is that they may and should include forms of subject matter local or provincial in character which it is not possible for standard tests to include. It is this characteristic that makes them of

such great help when they are used to supplement standard tests in the diagnosis of educational difficulties, since they enable a teacher so to extend the range of inquiry as to cover more nearly the pupils' range of achievement. Classroom tests are, of course, only as comprehensive as the teacher wishes to make them.

In general it may be said that standard tests are limited in their forms (that is, in the number of tests a teacher may have available for use with his pupils), but they are universal in the use of the results; whereas classroom tests are practically limitless in their forms, inasmuch as the teacher can make as many as he may please, but they are restricted in the use of the results to the class and to the pupils for whom they are made.

The classification of standard tests. It is when one attempts to classify standard tests themselves that it becomes difficult to find mutually exclusive bases for classification. In the following classification it is understood that there are many tests which might easily be classified (and properly are so classified) in more than one of the groups proposed. This is due partly to the fact that certain tests have a wide range of measurement, measuring several different forms of achievement at once, and partly to the fact that not enough is yet known of the elements in some fields to enable narrow measurements to be made. It is also due to the fact that, except theoretically, it is difficult to separate skills from either the knowledge which makes them possible or the attitudes which guide their expression.

Measurement of skills. In many respects the measurement of skills has been the least difficult problem with which the expert in measurement has had to deal. The school skills — reading, writing, spelling, arithmetic, and the like — have been so universally accepted as the foundation of elementary education, and the content of the work and the

expectations as to the results from them have such universal likeness, that they have formed the first and most successfully met problem in measuring. The school skills, also, because of the concreteness of their results in terms of achievement, are the most easily seen of all the primary objectives of school learning. Writing reveals itself in the ability to write, spelling in the ability to spell, and reading in the ability to read. These are concrete and readily judged evidences of learning — evidences far more easy to judge than are the elements of knowledge, of reasoning, or of attitudes and appreciations.

1. *Skills related to quality.* One of the measurements of skills is in terms of the degree of skill as evidenced by a given quality of product. As will be seen, this can be thought of in several ways, inasmuch as quality and quantity are inseparable, since any change in quality results in a change in the quantity of the quality. Quality as here used is understood to mean the evidence of degree of skill in terms of increasingly more artistic or æsthetic characteristics in the product, rather than the production of a greater amount of product of the same type. In handwriting, for example, a pupil may produce a sample of work which is no greater in amount of production than a second sample, yet the quality of the two differs, one being better than the other. It is even possible, but not necessarily so, that a sample which takes twice as long to produce as a second sample is of higher quality.

The tests that may be classified under this head are the direct quality scales, such as the handwriting scales and the composition scales. The measurement made by these scales is not in terms of quantity, but rather in terms of the quality of the product, the evidence of increasing skill being determined by the increase in the quality of the result. Fig. 4 shows a portion of the Thorndike Handwriting Scale, only qualities 6, 7, and 10 being illustrated.

2. *Skills related to time.* Many of the products of school instruction are closely bound up with the time that it takes to

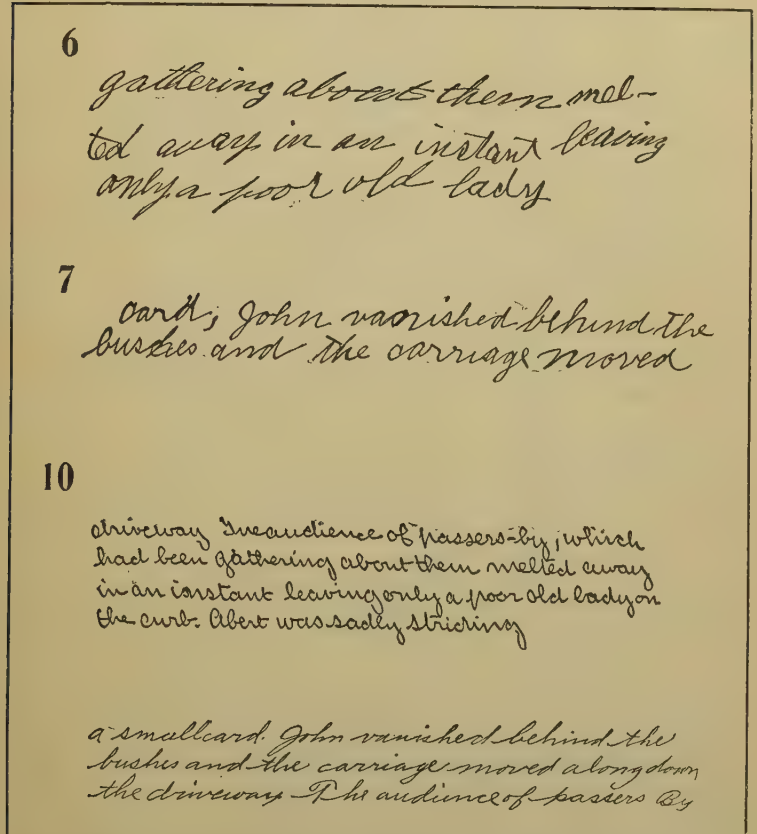

6
gathering about them melted away in an instant leaving only a poor old lady
7
card; John vanished behind the bushes and the carriage moved
10
driveway The audience of passers-by which had been gathering about them melted away in an instant leaving only a poor old lady on the curb. Albert was sadly sticking
a small card. John vanished behind the bushes and the carriage moved along down the driveway The audience of passers By

FIG. 4. Samples from the Thorndike Handwriting Scale[1]

accomplish them. This is particularly true of arithmetic achievements, where the efficiency in many cases is measured not so much in terms of the amount of product that the pupil

[1] Reproduced by permission of the Teachers College Bureau of Publications.

can do as in the accuracy of his work within given limits of time. Thus arithmetic tests are timed tests, except in cases of diagnostic tests involving other factors than time, such as tests to discover the ability to handle types of number combinations or types of computations. Fig. 5 shows part of a timed arithmetic test, the Courtis Standard Research Tests, Series B. Only a few of the problems are shown.

You will be given eight minutes to find the answers to as many of these addition examples as possible. Write the answers on this paper directly underneath the examples. You are not expected to be able to do them all. You will be marked for both speed and accuracy, but it more important to have your answers right than to try a great many examples.

927	297	136	486	384	176	277	837
379	925	340	765	477	783	445	882
756	473	988	524	881	697	682	959
837	983	386	140	266	200	594	603
924	315	353	812	679	366	481	118
110	661	904	466	241	851	778	781
854	794	547	355	796	535	849	756
965	177	192	834	850	323	157	222
344	124	439	567	733	229	953	525

FIG. 5. Courtis Standard Research Tests in Arithmetic. Test No. 1, Addition, Series B, Form 1[1]

3. *Skills related to difficulty.* A third measurement of skills, which in many cases is akin to the measurement of quality, is that involved in the measurement of the degree of difficulty in producing the result. This differs from the timed tests, where the effort is to measure the amount of the same type of product achieved in a given amount of time, in that the main effort is to measure the increasing difficulty of material with which an individual may cope. In this group are included many different types of tests. There are arithmetic tests where the degree of difficulty increases step by step. The

[1] Reproduced by permission of S. A. Courtis.

tests begin with simple problems (usually only one of any given degree of difficulty), and the problems become progressively more difficult to the end. In the reading tests which are based on difficulty the early selections are simple, and they progress to selections which are more difficult. The same progression from ease to difficulty is characteristic of all tests in this classification.

The selections in Fig. 6 are taken from the Thorndike-McCall Reading Scale, Form 2. They show the character of the easiest selection as compared with the most difficult. The test includes nine selections of graded difficulty between these two samples.

Another form of test now in use is called the spiral test. In this the type of exercise within the test is repeated in various sections of the test, becoming more difficult with each repetition. This form gives a wider possibility of diagnostic procedure, since it eliminates much of the danger of misinterpretation of single errors, a possible result from the narrower tests. The spiral plan of make-up is used in the Cleveland Survey Arithmetic Tests, a portion of which is shown in Fig. 7, on page 65. The illustration shows how the same type of element in the test increases in difficulty in successive repetitions.

Measurements of knowledge and knowledge values. Measurements of knowledge have been the main objective in the making and use of the traditional essay types of school examinations. As already explained, the defects in these tests have been the nonobjectiveness of the tests themselves and the restrictedness of their range. To overcome these difficulties the technics used in making the standard tests in skills (see Chapter V) have also been used in making standard tests of knowledge and knowledge values. These technics, which will be described in Chapter V, have to a great degree resulted in making the scoring of the tests more

Read this and then write the answers. Read it again if you need to.

Nell's mother went to the store on Water Street to buy ten pounds of sugar, a dozen eggs and a bag of salt. She paid a dollar in all. Nell and Joe went with her. On the way home on Pine Street, they saw a fire-engine with three horses.

1. Was the salt in a box or a bag or a can or a dish? _____

_____

2. How many eggs did she buy? _____

3. What did the children see on Pine Street? _____

_____

4. What street was the store on? _____

Read this and then write the answers. Read it again if you need to.

There are two methods by which one might make himself acquainted with anything made up of related parts; as, for example, a watch. He might take the watch apart, piece by piece, and while doing so study the details of its structure and the relation of its parts one to another. An operation like this, which begins with the whole and descends to the parts which compose the whole, is called analysis. The word means a taking apart or separating. Or he might begin with the parts, and, after some experiment and study, get an excellent knowledge of the watch by putting its parts properly together. An operation of this kind is called synthesis.

34. What two words are used in the paragraph to refer to both analysis and synthesis? _____

_____

35. Which two of these words have a meaning similar to *analysis*?

Draw a line under them.

method	separation	operation	assemble
related	experiment	resolution	dissemble

FIG. 6. Sample of the Thorndike-McCall Reading Test, Form 2[1]

[1] Reproduced by permission of Teachers College Bureau of Publications.

objective. They have also served in some degree to render them more comprehensive. As was pointed out, however,

SET A — Addition

1	6	9	0	4	1	7	9	3	2	1	3	6
2	6	5	1	2	3	7	6	0	4	5	8	9

SET E — Addition

5		2		9		2		6		1		4		9
2		8		8		8		3		4		6		7
2		8		0		5		4		2		5		1
0		5		7		0		8		5		3		5
4		1		6		6		8		4		4		3

SET J — Addition

7	9	4	7	2	9	6	7	7	8	9	4	3	2
5	2	5	1	9	6	9	1	8	0	5	3	1	1
4	4	8	9	4	2	6	5	5	7	3	7	7	6
2	8	1	4	8	4	7	1	4	1	4	7	6	6
6	2	4	3	5	7	0	4	1	8	6	0	9	1
0	7	8	2	1	1	4	6	8	5	2	2	6	8
5	5	5	8	5	3	3	5	2	1	3	9	3	6
1	3	1	5	2	9	7	3	1	3	9	5	4	9
8	6	3	2	4	2	1	3	3	7	2	6	5	7
3	1	9	7	3	3	6	7	9	4	2	3	4	5
2	4	6	7	6	8	0	6	8	9	8	4	2	2
9	8	3	1	7	5	6	1	4	4	5	8	9	2
9	8	5	9	6	5	6	7	5	4	6	8	9	4

FIG. 7. Sample of the Cleveland Survey Arithmetic Tests, Form 1[1]

standard tests must be built upon more or less universal forms of subject matter, and the forms of subject matter in which knowledge factors play the greatest part are extremely

[1] Reproduced by permission of the Public School Publishing Company.

variable as among different school systems. Geography and history, nature study or general science, English literature, and grammar vary greatly in different school systems both in content and placement within the curriculum. There is no uniformity among school systems, for example, as to where American history shall be taught, and there is less uniformity with respect to what in American history shall be taught. School systems which emphasize the junior high school model the curriculum of the first six grades differently from that of the systems where there is the four-year high school. Frequently rural and village schools have quite different emphases and content from those in city schools in the same district. All these factors tend not only to make difficult the construction of tests in the subject-matter fields but also to make it difficult to adjust them to the needs of different communities. Probably the greatest degree of success has been achieved in the field where knowledge is intimately related to skills, since here again the content is fairly universal.

1. *Knowledge as related to skills.* This classification is fairly closely associated with the entire group of skill tests discussed in a previous section. There is in this classification, however, a slightly different emphasis, although in some cases the subject matter involved is similar. These tests place a larger degree of emphasis on the knowledge or knowledge values necessary to complete the test than on the essential degree of skill required. For this reason the important group of diagnostic tests which have developed are here included, as well as tests involving knowledge as an evidence of degree of skill.

In general the spelling tests place a slightly higher emphasis on the knowledge of spelling than they do on skill in spelling, as is shown by the fact that these tests are rarely timed tests, but are rather dictated at a speed which the

teacher may feel is best for the ability of the class. The final rating is given in terms of knowledge of the way to spell words correctly rather than in terms of the time taken to complete the test. A test of this type is shown below.

LIST 4

1. she	Is *she* at home?	she
2. man	That *man* is good	man
3. good	Be a *good* boy	good
4. this	*This* book is mine	this
5. over	The ball is *over* the fence	over
6. say	You must *say* your prayers	say
7. belong	Does this pen *belong* to you?	belong
8. stand	Please *stand* up	stand
9. sick	The doctor is *sick*	sick
10. foot	I hurt my *foot*	foot
11. sister	My *sister* is here	sister
12. inside	Come *inside* the tent	inside
13. mail	Please *mail* my letter	mail
14. two	I have *two* apples	two
15. began	I *began* work last week	began
16. watch	*Watch* that man	watch
17. happen	It does not *happen* often	happen
18. sight	I lost *sight* of the bird	sight
19. bridge	The *bridge* is large	bridge
20. front	The *front* of the car is smashed	front
21. elect	I hope they *elect* my choice	elect
22. refuse	I *refuse* to go	refuse
23. pleasure	It was a great *pleasure*	pleasure
24. vacation	I had two months' *vacation*	vacation
25. repair	I must *repair* my wheel	repair

FIG. 8. Sample of the Morrison-McCall Spelling Scale, List 4[1]

Fig. 8 shows part of one of eight different forms of the Morrison-McCall Spelling Scale.

[1] Reproduced by permission of the World Book Company.

The diagnostic forms of tests, which are to be found for various phases of English grammar, arithmetic, language, reading, and the like, are designed to make possible, by means of an analysis of the results of the tests, a discovery of the types of errors which have been made, with a view to correcting them. Such tests place greater emphasis upon the knowledge which leads to a skill than upon the skill itself, as is illustrated particularly in the diagnostic tests in arithmetic and reading.

In Fig. 9 will be found a part of a diagnostic test in silent reading as published by Gates. It forms one of the large number of tests devised by Dr. Gates for the diagnosis of special difficulties in the reading process.

As an example of the emphasis placed upon knowledge rather than skill, a diagnostic test in arithmetic by Monroe is shown in Fig. 10. This test covers the knowledge of pupils in handling integers in Grades 4–8 inclusive, and is so arranged that individual errors can be located as types for individual study.

2. *Knowledge as related to reasoning.* A second classification in which knowledge is emphasized consists of a large group of tests in which the use of knowledge in reasoning is the main emphasis. In this group are included tests in the general fields of knowledge, such as geography, history, and the like. The great difficulty in the making and use of these tests is the difference in content and emphasis between school systems and even between classes within the same school system. Where the beginnings of American history may be taught in the fourth grade in one system, they may be delayed until the sixth grade in another. A certain form of geography may be taught in one school system in the fifth grade, and the same form may not be considered at all in another. The differences in curriculum make the construction of tests in these fields most difficult and also make

1. Little Betty liked to have long rides on her sled. She was too little to go by herself so her father played horse and pulled her along. One day Betty's father was pulling her along a very bumpy road. He was looking straight ahead, when suddenly he felt that the sled was very light. He looked back. Betty had fallen off the sled!

Betty's father ran away
Betty's father put her on the sled
Betty stayed there all day
Betty put her father on the sled

2. Tom jumped up on the cow's broad back. "Get up, Bossy!" he shouted. The cow did get up and so quickly Tom almost fell off her back. Then she started off. Bump, bump went Tom. "Whoa," he cried for he did not like the ride. But the cow would not stop. As they reached the pasture, Tom was thrown to the ground.

Tom rode back on the horse
He said, "Old Bossy, I'll not ride you again"
Tom gave the good cow some hay
Tom jumped down from the cow's back

23. A little chick was out looking for some worms one day. Along came a pretty yellow caterpillar. He was so bright and pretty that the little chick could not help eating him. But oh, what a bitter taste he had. It made the little chick feel quite sick. After a while he was able to eat again. Along came another pretty yellow caterpillar.

The caterpillar ate the chick
The chick ate the caterpillar
The chick ate the turtle
The chick did not eat the caterpillar

24. There are a number of plants which catch insects for food. Some of them seem cruel. There is one plant called the Venus flytrap. It has leaves that grow in pairs which will quickly snap together if one of them is touched. One day I was watching one of these plants. A fly flew down and sat on one of the trap leaves.

I ate the fly
The leaves caught the fly
My finger was caught in the leaf
The fly rested and flew away

FIG. 9. Sample of the Gates Silent Reading Test, Type B (Reading to predict the Outcome of Given Events), Form 1, Grades 3–8[1]

[1] Reproduced by permission of Teachers College Bureau of Publications.

Test 1 — ADDITION

At. _____

Rt. _____

4	5	2	0	1	7	6	7	3	2	3	9
7	5	6	3	1	2	8	7	8	4	3	4
2	9	7	8	4	3	4	0	9	0	6	5

Test 2 — SUBTRACTION

At. _____

Rt. _____

37	94	60	27	39	41	77	53
5	8	3	6	7	8	3	9

Test 3 — MULTIPLICATION

At. _____

Rt. _____

6572	6750	5863	3754	2845
6	9	2	5	8

Test 4 — DIVISION

At. _____

Rt. _____

8)3840 4)7432 7)2534 3)8430 6)4680

FIG. 10. Sample of the Monroe Diagnostic Tests in Arithmetic. Operations with Integers, Part I, Tests 1–4[1]

[1] Reproduced by permission of the Public School Publishing Company.

THE LOCATION OF CITIES. The cities are represented on the map by little crosses with arrows leading out to a figure on the map. For example, Los Angeles is number 11. Put the right number after each of these cities. If the cities are not found on the hemisphere maps at the top of the pages look on the enlarged maps below:

CITIES	PUT YOUR ANSWERS HERE
Los Angeles	11
Chicago	
New York	
Liverpool	
Buenos Aires	
Detroit	
Pekin	
Philadelphia	
San Francisco	
Paris	

FIG. 11. Portion of the Gregory-Haggerty Geography Test, Form B, Part 2, on Locational Geography and Map Study[1]

the interpretation of the results somewhat misleading. As a result there are two general ways taken to make such tests. The first is to divide the subject into its logical divisions. In geography this would mean to publish one test on locational

[1] Reproduced by permission of C. A. Gregory, University of Cincinnati, Ohio.

geography and map study, another on climate and rainfall, and another on plants, animals, people, and the like. The second is to give tests on separate units, such as the United States, Europe, Asia, South America, and so on. The first type lends itself to reasoning within the test; the second, to factual testing. In Fig. 11 is shown the first type, which is a portion of the Gregory-Haggerty Geography Test, Form B, Part 2, on Locational Geography and Map Study.

A sample of the second type, Fig. 12, shows several small portions of the test on the United States taken from the Buckingham-Stevenson Place Geography Tests.

Say to the pupils, **I am going to give you some more exercises like these.** (Read the following statement. Repeat once so that all will understand. **Name the state in which each of the following is located:** Read the number of the exercise first; as, number 6, San Francisco; number 7, Buffalo; etc. Allow enough time for all the pupils to answer. Tell the pupils that the city referred to is the largest city of that name).

FORM 1

6. San Francisco
7. Buffalo
8. Nashville
9. New Orleans
10. Cleveland
11. Toledo
12. Dallas

Name the State in which Each of the Following is Located

26. Mt. Shasta
27. Mt. Washington
28. Mohave Desert
29. Yosemite Falls

Name the State which is Next West of Each of the Following States

30. New Hampshire
31. Alabama
32. Indiana
33. North Dakota

Name the State which is Next North of Each of the Following States

36. Pennsylvania
37. South Dakota
38. South Carolina
39. Missouri

FIG. 12. Buckingham-Stevenson Place Geography Tests, United States, Form 1.[1] (Small portions only are shown)

[1] Reproduced by permission of the Public School Publishing Company. Copyright, 1922.

In arithmetic the tests which belong in this classification are known as reasoning tests. The first of these to be made was the Stone Reasoning Test in Arithmetic (see Chapter III), part of which is reproduced in Fig. 13.

Solve as many of the following problems as you have time for; work them in order as numbered.

1. If you buy 2 tablets at 7 cents each and a book for 65 cents, how much change should you receive from a two dollar bill?

2. John sold 4 Saturday Evening Posts at 5 cents each. He kept $\frac{1}{2}$ the money and with the other $\frac{1}{2}$ he bought Sunday papers at 2 cents each. How many did he buy?

3. If James had 4 times as much money as George, he would have $16. How much money has George?

4. How many pencils can you buy for 50 cents at the rate of 2 for 5 cents?

FIG. 13. Sample of the Stone Reasoning Tests, Test 1[1]

Tests of ideals or appreciations. This field of learning, although considered by teachers as one of their great objectives in teaching, has been most difficult to define and measure. One reason for the difficulty experienced is that these objectives are either disputed in their elements or else are intangible. Attempts have been made in certain fields to locate the elements which are tangible and to construct tests to measure them. One of the more interesting cases of this type was made by Voelker. A sample of it is reproduced in Fig. 14.

The greatest success in this field has been in connection with appreciations in English literature and in poetry, where, as one might expect, the greatest effort has been placed. A sample of this type of test is shown in Fig. 15, which is taken from the Logassa and McCoy-Wright Test for Appreciation of Literature, Test 3, Reaction to Sensory Images. Only two small portions of the test are shown, but they illustrate the method used.

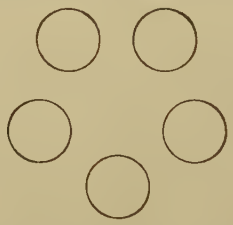
[1] Reproduced by permission of Teachers College Bureau of Publications.

Can the subject be trusted not to peep when he is placed on his honor to keep his eyes closed?

DIRECTIONS: The principle and the method in this test are the same as in the Profile Test described under Series 1. The apparatus only is different.

Each subject is handed a card 12 inches by 12 inches. On the card are five circles each five-eighths of an inch in diameter. The circles are located at the angles of an imaginary equilateral pentagon, the sides of which are six and one half inches long.

Following is a diagram of the arrangement of the circles, the dimensions of the imaginary pentagon being reduced in size.



The subject is instructed to look carefully at the circles and then with eyes closed to try to place a pencil mark in each of the five circles as his hand sweeps around the circumference of the pentagon. If he places a mark in each circle as he goes once around, he writes on a card, "First trial, correct"; if he fails, he writes, "First trial, wrong." He is given five trials. The individual reporting the highest number of successes is awarded a prize.

SCORING: A subject who reports one or more successes in the five trials is scored 0. A subject who reports only failures is scored 10. This feat is impossible of accomplishment as far as it has been possible to determine. The examiner himself, after long practice and fifty successive trials, was unable to score a single success. No trustworthy subject has been found who is able to accomplish the feat.

FIG. 14. The Voelker Honesty Tests, Series 2, No. 9[1]

[1] Reproduced by permission of Paul F. Voelker from pages 108–109 of "The Function of Ideals and Attitudes in Social Education," Teachers College Contributions to Education, No. 112, Teachers College, New York, 1921.

3.	The ship began to dip into the southwest swell, and the softly luminous sky of low latitudes took on a harder sheen from day to day above their heads: it arched high above the ship, vibrating and pale, like an immense dome of steel, resonant with the deep voice of freshening gales. The sunshine gleamed cold on the white curls of black waves. Before the strong breath of westerly squalls the ship, with reduced sail, lay slowly over, obstinate and yielding. She drove to and fro in the unceasing endeavor to fight her way through the invisible violence of the winds; she pitched headlong into dark smooth hollows; she struggled upwards over the snowy ridge of great running seas; she rolled, restless, from side to side, like a thing in pain.	sight sound touch temperature smell taste muscle
4.	Why does the sea moan evermore? Shut out from heaven it makes its moan, It frets against the boundary shore: All earth's full rivers cannot fill The sea, that drinking thirsteth still.	sight sound touch temperature smell taste muscle

FIG. 15. Samples of the Logassa and McCoy-Wright Tests for Appreciation of Literature. Test 3, Reaction to Sensory Images.[1] (Directions ask the pupil to check the senses etc. to which the author appeals)

Fig. 16, a part of a test by Abbott and Trabue, in the field of appreciation of poetry, shows another attempt which is successful in measuring a form of appreciation on the part of pupils.

It may be confidently expected that if the present trend in education continues (and it is probable that the æsthetic, appreciative, and idealistic results of education will continue to be stressed), more emphasis will be placed on these elements in testing. Therefore the next few years should see greater discussion of this phase of testing and more attempts to measure the results.

Tests of ability, capability, potential ability, or intelligence. The development of this type of test has been discussed in Chapter III. These tests, commonly called intelligence tests,

[1] Reproduced by permission of the Public School Publishing Company.

Read the poems A, B, C, D, trying to think how they would sound if read aloud. Write "Best" on the dotted line above the one you like best as poetry. Write "Worst" above the one you like least.

Set 4. Potatoes

A (. . . . .)

When you have to eat potatoes
With other things,
And their jackets taken off,
That might be good for Dukes and Kings.
But if you want to know how they taste
The way Nature meant you should,
Cook them over a fire of rubbish
And eat them in the woods.

B (. . . . .)

Potatoes on the table
To eat with other things,
Potatoes with their jackets off,
May do for Dukes and Kings.
But if you wish to taste them
As Nature meant you should,
Why, cook them at a rubbish fire
And eat them in a wood.

C (. . . . .)

Potatoes at the table
To eat with other food,
Potatoes carefully prepared,
May be very good.
But if you want to eat them
So that you'll want some more,
Cook them over an open fire
And eat them out of doors.

D (. . . . .)

Ho, for the Hot Potato!
Of all the feast the king!
I love it more than all the rest,
Ice-cream or anything!
And of all the treasures of Nature's best
The one I most admire
Is the glowing hot Potato,
Just coming from the fire.

FIG. 16. Sample of the Abbott-Trabue Exercises in Judging Poetry, Series X, Set 4[1]

[1] Reproduced by permission of Teachers College Bureau of Publications.

may be classified in several ways, but they are probably most usefully classified according to their form.

1. *Individual intelligence tests.* An individual intelligence test is one which must be administered in such a manner as to enable the examiner to give his undivided attention to the person being examined. The Stanford Revision of the Binet-Simon Test is one of this type. The examiner can give the test to only one child at a time. This test is one which requires a high degree of examining skill, and results should not be used until the examiner has had sufficient practice to be able to discriminate readily among the answers and in the scoring.

A part of this test is given in Fig. 17, together with the directions for giving the test and scoring the answers.

There are several other revisions of the Binet-Simon Scale. They have the same general characteristics as the Stanford Revision, but are designed either to make the administration of the test somewhat easier or to make the scoring less subjective.

The Detroit Kindergarten Test is an individual intelligence, or ability, examination. It is designed for use in testing the reactions of pupils entering the kindergarten, and can therefore be used for purposes of classifying such young children. A portion of the directions, the body of the test, and the scoring sheet for the test is given in Fig. 18.

2. *Group intelligence tests.* As has been pointed out, the group tests for the measurement of intelligence have a far more widespread value for the measurement of school children than have the individual tests. A group test is one that can be administered to an entire class of pupils at one time, the pupils receiving directions together and recording their answers individually. As a result the examiner can do the scoring and can make his judgments as to the quality of the answers at some later time.

YEAR V. (6 tests, 2 months each, or 4 tests, 3 months each.)

*1. Comparison of weights. (2 of 3. E. may illustrate procedure.)

3–15_____ 15–3_____ 3–15_____

*2. Colors. (No error.)

Red_____ Yellow_____ Blue_____ Green_____

*3. Æsthetic comparison. (No error.)

Upper pair_____Middle_____Lower_____

4. Definitions, use or better. (4 of 6.)

Chair _____ Doll_____

Horse _____ Pencil_____

Fork_____ Table _____

1. ***Comparison of Weights***

Place the 3 and 15 gram weights before S., 2 or 3 inches apart. Say, "You see these blocks. They look just alike, but one of them is heavy and one is light. Try them and tell me which one is heavier." Repeat instructions if necessary, saying, "Tell me which one is the heaviest." If S. merely points without lifting blocks, or picks up one at random, say, "No, that is not the way. You must take the blocks in your hands and try them, like this." (Illustrate.) Give second trial with position of weights reversed; third trial with weights in same position as first.

Credit if two of three comparisons are correct.

2. ***Naming Colors***

Show card (V 2) and say, pointing to colors in the order, red, yellow, blue, green, "What is the name of that color?"

Credit if all colors are correctly named, without marked uncertainty.

3. ***Æsthetic Comparison***

Show pairs of faces in order from top to bottom of card (V 3). Say, "Which of these two pictures is the prettiest?"

Credit if all *three* comparisons are made correctly.

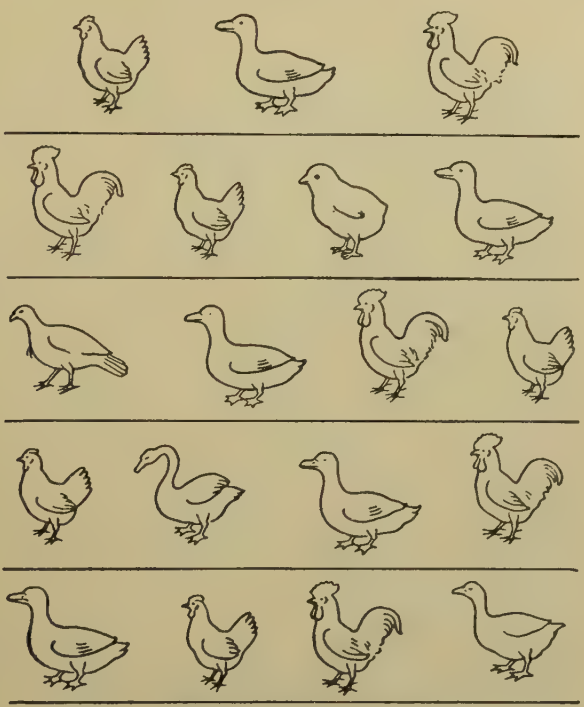
4. ***Definitions: Use or Better***

Say, "You have seen a chair. You know what a chair is. Tell me, what is a chair?" If necessary urge as follows: "I am sure you know what a chair is. You have seen a chair." "Now, tell me, what is a chair?" If S. rambles say, "Yes, but tell me; what is a chair?" Same for horse, fork, doll, pencil, table.

Credit if four words out of the six are defined in terms of use or better. (See *The Measurement of Intelligence*, p. 168.)

FIG. 17. Sample of the Stanford Revision of the Binet-Simon Scale for the Measurement of Intelligence, Tests 1–4, fifth year[1]

[1] Reproduced by permission of Houghton Mifflin Company.



TEST PROCEDURE

The boldface type shows the directions that the teacher gives the pupil; remarks to the teacher are in parentheses.

Test 2

(Point to top row each time.) **See these things here. Now show me the one here that is not up there.** (Repeat for each row.)

The maximum score on the test is 30 points. One point credit is allowed for each row, except in Test 7, where one point is given for each item. No credit is allowed unless the correct item or items *only* are marked.

Test	Item or Items Correct	Points
2	Chick, turkey, swan, goose	4

FIG. 18. Sample of the Detroit Kindergarten Test, Test 2[1]

[1] Reproduced by permission of the World Book Company.

There are two types of group tests: the non-language tests and the language tests. The non-language tests are as a rule designed for the measurement of intelligence of children in the primary grades before they have developed much facility in reading; the language tests are designed for those who can read and write.

One of the non-language tests, the Pressey Primary Classification Test (part of which is reproduced in Fig. 19, together with the accompanying directions and scoring key), determines one type of ability by testing skill in the discovery of absurdities. This test is intended for use only in primary grades. Since intelligence (in its development at least) is largely independent of learning, such tests as these are not suitable for testing adult illiterates.

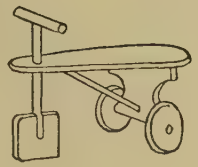
A second primary test for intelligence is the Pintner-Cunningham Primary Mental Test, part of which is reproduced in Fig. 20, on page 82. This test is suitable for use through the second grade.

Of the language forms of intelligence tests there are several kinds, designed for different purposes — for use in different grades and for measuring different types of ability. One of these tests, the National Intelligence Test (commonly known as the N.I.T.), has already been mentioned. This test was modeled after the Army Alpha, the intelligence test used so widely with soldiers during the World War. It consists of a number of types of work, each of them timed, and is designed to measure different types of reaction. Part of two sections of this test will serve to show its character (see Fig. 21).

The Haggerty, the Dearborn, and the Otis group intelligence tests are of somewhat the same general character and are reliable tests. A different type of group intelligence test has been developed by McCall and some of his students. This scale, known as the McCall Multi-Mental Scale, is reproduced in part in Fig. 22, on page 84.

Now put your pencils down on your desks. In each of the other pictures on the page there is something wrong. Find the part that is wrong, and cross it out. Pick up your pencils and start. Finish all the pictures on the page.

(Time 2 min. Do not help the children but see that they keep working.)

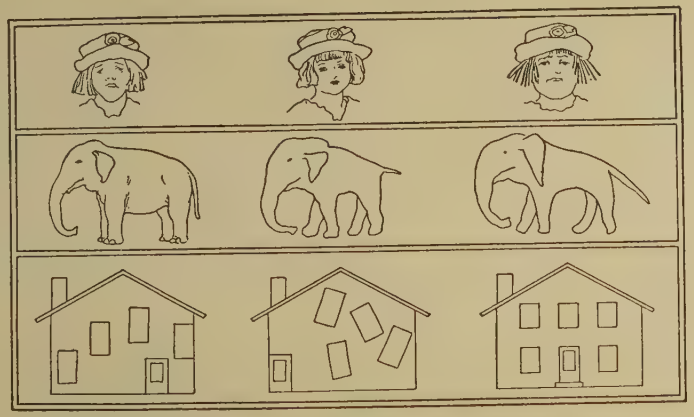


Directions for Scoring

Each item correctly marked counts 1 point. To be scored correct **one, and only one,** part may be marked. If more than one part is marked, count that item zero.

FIG. 19. Sample of the Primary Classification Test, by L. W. Pressey[1]

[1] Reproduced by permission of the Public School Publishing Company.



Test 2. Æsthetic Differences (pages 4 and 5)

Page 4. **This is something different. Look at the three little girls at the top of the page.** (Point.) **I want you to find the prettiest. Don't tell any one, but look at all the girls. Then put a mark on the prettiest girl. Go ahead!** (Wait until all have finished.) **Now look at all the elephants. Mark the prettiest elephant.** (Pause not longer than 10 seconds.) **Look at all the houses. Mark the prettiest house.** (Pause 10 seconds.)

FIG. 20. Sample of the Pintner-Cunningham Primary Mental Test, p. 4, Test 2[1]

[1] Reproduced by permission of the World Book Company.

Test 1

Find all the answers as quickly as you can.
Write the answers on the dotted lines.
Use the sides or bottom of the page to figure on.

Begin here

1 One trolley ride costs 5 cents. How much do seven trolley rides cost? *Answer* ____

2 The boys rode 21 miles and walked 4 miles more. How far did they go in all? *Answer* ____

3 How many cents does it cost for a 5-cent stamp and a 2-cent stamp? *Answer* ____

4 Thirty-two children march 4 in a row. How many rows are there? *Answer* ____

Test 3

In each row draw a line under each of the two words that tell what the thing always has.

SAMPLES
- **man** (body cane head shoes teeth)
- **dog** (blanket chain collar legs nose)
- **house** (cellar paint room servants walls)

Begin here

1 **father** (bones forehead mustache son watch)
2 **child** (back book clothes hat mouth)
3 **sandwich** (bread cheese eggs filling ham)
4 **sheep** (barn bell pasture tail wool)

5 **book** (music pages pictures printing stories)
6 **telephone** (desk directory mouthpiece operator receiver)
7 **shower** (clouds mud rainbow raindrops umbrellas)
8 **teapot** (body heat spout tea water)

FIG. 21. Sample of the National Intelligence Test, Scale A, Form 3[1]

[1] Reproduced by permission of the World Book Company.

"Pencils down! . . . . Listen carefully. In the same way, draw a line through the one and only one word which does not belong with the others in each of the sets on this page. Do set number 1 first, then set number 2, and so on. Be sure not to skip any set. Mark some word in it before passing on to the next set. You will have plenty of time, if you don't waste it. If you finish before time is called, go back and make sure you have made no mistakes. If you decide to change your mark, don't forget to erase your first mark. Now once more, to make certain you understand every point. Be sure to draw a line right through the one and only word in each set that does not go with the others. Be sure to do set 1 first, then set 2, and so on without skipping any set. Be sure to use every minute of the time, that we may be proud of the score for our school. Ready with pencil! Go!"

1	2	3	4	5
fly	cup	horse	lesson	grass
burn	fork	calf	problem	coal
gasoline	saucer	colt	teacher	carbon
coal	bowl	hen	learn	tar
wood	knife	cow	solve	soot
6	**7**	**8**	**9**	**10**
robin	high	irrigate	black	word
geranium	low	land	hot	paragraph
elephant	cat	soil	white	sentence
poppy	fever	cultivate	star	style
bluebird	dangerous	navigate	cold	composition

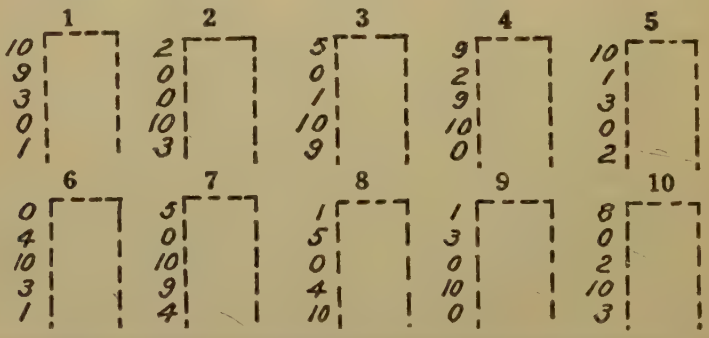


FIG. 22. Sample of the McCall Multi-Mental Test and Key[1]

[1] Reproduced by permission of the Teachers College Bureau of Publications.

In addition to these forms of group intelligence tests there are a number of tests designed especially for rating pupils in the junior and senior high schools. Such tests are frequently tests of mental ability. Two examples are the Miller Mental Ability Tests and the Otis Self-administering Tests of Mental Ability, Higher Examination. A part of the latter is given in Fig. 23 to show how it differs from other tests.

EXAMINATION BEGINS HERE:

1. The opposite of hate is (?)
   1 enemy, 2 fear, 3 love, 4 friend, 5 joy . . . . . ( )
2. If 3 pencils cost 5 cents, how many pencils can be bought for 50 cents?. ( )
3. A bird does not always have (?)
   1 wings, 2 eyes, 3 feet, 4 a nest, 5 a bill . . . . . ( )
4. The opposite of honor is (?)
   1 glory, 2 disgrace, 3 cowardice, 4 fear, 5 defeat . . . . . ( )
5. A fox most resembles a (?)
   1 wolf, 2 goat, 3 pig, 4 tiger, 5 cat . . . . . ( )
6. Quiet is related to sound in the same way that darkness is related to (?)
   1 a cellar, 2 sunlight, 3 noise, 4 stillness, 5 loud . . . . . ( )

FIG. 23. Sample of the Otis Self-administering Tests of Mental Ability, Form A, Higher Examination[1]

Practice tests. An important group of tests which is now much in use is known as the practice test. These tests were originated by Dr. Courtis in Detroit as a regular part of his teaching plan with Detroit teachers. Others have adapted the idea, and there are a number of such tests on the market.

The Courtis Standard Practice Tests in Arithmetic are not tests in the strictest sense, although there are tests in the materials. They consist of forty-eight graded lessons, each of which contains practice material in some phase of the arithmetical operations involving whole numbers. These are printed on cards in sufficient quantity for each pupil to have a complete set. In addition to these cards there are

[1] Reproduced by permission of the World Book Company.

research tests for checking the development of the pupils. Each pupil can practice an operation according to his needs. He places one of the cards under a tissue sheet so that he can read the printing on the card through the sheet. He solves the problems that are presented and then checks his answers himself. Part of a Courtis test is reproduced in Fig. 24.

Lesson No. _____ Form _____ Date _____

Name _____ Grade _____

90472	66119	81185	153776
50196	52134	37984	67017
5			
80836	66194	121961	84097
49178	36490	90492	26817
		10	
51462	132577	97182	
26861	53039	20702	33 A
89729	76514	111694	
30606	27811	55475	
15			
99484	94093	11043	
69924	59383	2748	

SCORES TRIAL

Tried ___ No. ___

Right ___ Form A

FIG. 24. Sample of the Courtis Standard Practice Tests in Arithmetic, Lesson No. 33, Subtraction[1]

[1] Copyright by the World Book Company. Reproduced by permission.

Another example of a practice test is the McCall-Crabbs Standard Test Lessons in Reading (see Fig. 25).

TEST LESSON 15

Haven't you often wondered what our schools were like in the days of long ago? The first schools were often held in the homes of one of the women of the village. The children all gathered in one room, which was usually the kitchen. The teacher kept the naughty ones from straying by tying them to her chair. Later, one-room schoolhouses were built. At one end of the room was a wide fireplace. The windows were made of oiled paper instead of glass. There were long benches made of logs split in two, running across the room. It was not until about one hundred and twenty-five years ago that schoolhouses with an upstairs were built.

1. In our first schools the children gathered in the (*a*) dining room; (*b*) living room; (*c*) kitchen; (*d*) on the porch.
2. These schools were taught by (*a*) men; (*b*) women; (*c*) girls; (*d*) boys.
3. The next schools built contained (*a*) two; (*b*) three; (*c*) one; (*d*) four rooms.
4. Our story said the teacher (*a*) spanked; (*b*) praised; (*c*) scolded; (*d*) tied the naughty children.
5. Our first schools were held in (*a*) barns; (*b*) homes; (*c*) stores; (*d*) forts.
6. At one end of the room was a (*a*) stove; (*b*) chair; (*c*) bench; (*d*) fireplace.
7. The windows were made of (*a*) boards; (*b*) glass; (*c*) paper; (*d*) skin.
8. Were the benches made of (*a*) boards; (*b*) iron; (*c*) stone; (*d*) logs? (*Item 9 is omitted.*)

No. right	0	1	2	3	4	5	6	7	8	9	10
G score	1.6	1.8	2.7	3.1	3.5	3.8	4.1	4.4	4.9	5.9	...

FIG. 25. Sample of the McCall-Crabbs Standard Test Lessons in Reading, Book Two, Test Lesson 15[1]

[1] Reproduced by permission of the Teachers College Bureau of Publications.

Surveys and batteries. Because of the increasing use of groups of tests for purposes of survey and classification, a number of tests have been devised to serve that purpose. These survey tests are in reality batteries of tests covering several fields. They are group tests, and can be given to pupils at one sitting, in the same way as any other group test might be given. Two of these tests are the Stanford Achievement Test and the Illinois Survey.

The Stanford Achievement Test consists of tests in arithmetic, reading, science, general information, history, and literature. Parts of the test are also published separately. Fig. 26 illustrates the form in which they are presented.

The Illinois Survey consists of several tests, including special tests in reading and arithmetic. It is somewhat similar in character to the Stanford Achievement Test. Fig. 27 gives an illustration of its character.

Another type of survey or classification test combines tests of intelligence with educational tests in one battery, with the idea of classifying pupils on the basis of a single test. Such a test is the Otis Classification Test, which consists of an achievement test covering reading, spelling, language and grammar, arithmetic, geography, history and civics, hygiene, vocabulary, music and art, together with a test devoted to the measurement of intelligence or ability. Part of this test is reproduced in Fig. 28, on page 91.

Chapter summary. Tests may be classified either as standard tests or as classroom tests. Standard tests are characterized by the care exercised in their making, by the objectivity of their scoring, by the comprehensiveness within the field measured, by the universal character of their content, and by the fact that they are accompanied by norms, or standards of achievement. Classroom tests are usually made by the classroom teacher for his own class. They are more objective than the traditional type of essay examination, but are less

TEST 1. READING: PARAGRAPH MEANING

Sample: Dick and Tom were playing ball in the field. Dick was throwing the ball and ____ was trying to catch it.

Write JUST ONE WORD on each dotted line.

1 Fanny has a little red hen. Every day the hen goes to her nest and lays an egg for Fanny to eat. Then she makes a funny noise to tell Fanny to come and get the ____.

2 A kitten can climb a tree, but a dog cannot. This is very lucky for Nellie's kitten. Every time Joe's big dog comes along the kitten climbs a tree and the ____ cannot follow.

3 Anna had never seen a squirrel in her life, although she had always wanted to very much. One day when she was playing under a tree she heard a funny little noise over her head. She looked up, and what do you think she saw? Up there in the _____ was the very thing she had always wanted to see, a _____.

TEST 5. ARITHMETIC: REASONING

Find all the answers as quickly as you can.
Write the answers on the dotted lines.
Use the blank sheets of paper to figure on.

Begin here.

1 How many are 3 eggs and 2 eggs? *Answer* _____

2 Mary is 7 years old. How old will she be in 3 years? *Answer* _____

3 A hen had 9 chicks and 3 of them died. How many were left? *Answer* _____

4 Milk costs 8 cents a pint and the milkman is going to raise the price 2 cents. What will it then cost? *Answer* _____

5 If you buy a pencil for 4 cents and pay for it with a dime, how much change should you get? *Answer* _____

TEST 8. LANGUAGE USAGE

Samples

Apples $\frac{\text{is}}{\underline{\text{are}}}$ good.

He $\frac{\underline{\text{told}}}{\text{telled}}$ me.

1 I $\frac{\text{calculate}}{\text{expect}}$ to go soon.

2 Last year uncle $\frac{\text{gave}}{\text{give}}$ me a pair of skates.

3 His leg was $\frac{\text{broke.}}{\text{broken.}}$

4 They have $\frac{\text{gone}}{\text{went}}$ to town.

5 He isn't $\frac{\text{any}}{\text{no}}$ better than you.

6 Always $\frac{\text{bathe}}{\text{wash}}$ your hands before eating.

FIG. 26. Samples of the Stanford Achievement Test[1]

[1] Reproduced by permission of the World Book Company.

Test No. 1 — ANALOGIES

No. Right _____

1	eat — bread:: drink — water iron lead stones _____	1
2	finger — hand:: toe — box foot doll coat _____	2
3	shoe — foot:: hat — kitten head knife penny _____	3
4	dress — women:: feathers — bird neck feet bill _____	4
5	dog — puppy:: cat — kitten dog tiger house _____	5

Test No. 5. — VERBAL INGENUITY

No. Right _____

1 the cat at see.

2 boy was sky the sick.

3 Bread sweep will the kitchen I.

4 are going yesterday to-morrow we.

5 me mine give my straw hat.

Test No. 6 — ARITHMETICAL INGENUITY

No. Right _____

1 2 3 9 4 5

2 4 6 7 8

9 8 7 6 5 2

11 10 8 6 4 2

5 7 10 15 20 25

FIG. 27. Samples of the Illinois Examination, Form 1, General Intelligence [1]

[1] Reproduced by permission of the Public School Publishing Company.

100. Which one of these states is the most level?
1 California, 2 Colorado, 3 Kansas, 4 Pennsylvania, 5 Montana ( )

101. When it is 12 o'clock at Chicago, what time is it at Denver? . . . . . ( )

102. The Spanish-American War was fought during the administration of (?)
1 Cleveland, 2 McKinley, 3 Roosevelt, 4 Taft, 5 Wilson . . . ( )

103. Which one of the following men was prominent in the Civil War period?
1 James Monroe, 2 Alexander Hamilton, 3 Jefferson Davis, 4 Grover Cleveland, 5 John Paul Jones . . . . . ( )

104. The judges of the Supreme Court of the United States are (?)
1 appointed by the President for office during his term, 2 elected by the people every four years, 3 chosen by Congress every four years, 4 appointed by Congress for life, 5 appointed by the President for life . ( )

105. The tympanic membrane is part of the (?)
1 mouth, 2 nose, 3 eye, 4 ear, 5 throat . . . . . ( )

FIG. 28. Sample of the Otis Classification Test, Form B, for Grades 4–8[1]

objective than standard tests. They may include subject matter which is of local, or provincial, type, but they are lacking in norms, or standards, which would make it possible to use them for comparison.

Standard tests are difficult to classify in mutually exclusive groups. They are here classified according to the following schematic outline:

1. Measurements of skills
   a. Skills as related to quality
      Handwriting and composition scales
   b. Skills as related to time
      Timed arithmetic tests
   c. Skills as related to difficulty
      Tests in arithmetic, spelling, reading, and the like where there is a gradation in the material; also includes spiral tests
2. Measurements of knowledge and knowledge values
   a. Knowledge as related to skills
      Spelling tests, and diagnostic tests of many kinds in reading, arithmetic, English, and the like

[1] Reproduced by permission of the World Book Company.

b. Knowledge as related to reasoning
Reasoning tests in arithmetic; also includes most tests in geography, history, and the like

3. Measurements of ideals or appreciations
Includes tests in the appreciation of literature and poetry as well as tests of honesty and character

4. Tests of ability, capability, potential ability, or intelligence
a. Individual intelligence tests
The basic test is the Stanford Revision of the Binet-Simon Scale; other revisions are also included, together with other individual intelligence tests
b. Group intelligence tests
(1) Non-language tests
Includes tests largely designed for the primary or pre-school units; also includes tests for adult illiterates
(2) Language tests
Designed for literates throughout the educational range

5. Practice tests
Combinations of teaching materials and test materials put into convenient form for the use of the teacher or pupil

6. Batteries of tests, or test surveys
a. Achievement tests
Combinations of the foregoing types of achievement tests arranged for purposes of classification
b. Achievement and mental tests combined
A combination of both types of tests for the purpose of bringing about a classification of children through one test period

SELECTED BIBLIOGRAPHY

GREENE, H. A., and JORGENSEN, A. N. The Use and Interpretation of Educational Tests, chap. ii. Longmans, Green & Co., New York, 1929.
RUSSELL, C. Classroom Tests, chap. ii. Ginn and Company, 1926.

CHAPTER V

THE CONSTRUCTION OF TESTS

The mechanical construction of tests. As tests are now published, practically every element that might have some effect on the form of the results, or that might in some way affect them, has been mechanized. When the tests are distributed, all the materials which a teacher might need are provided, either on the test sheets or in the manuals. Before making any attempt to give a test all the materials should be carefully examined, the purpose of each understood, and the significant details noted. Some of the details of the mechanical construction of tests are simple both in appearance and in interpretation; others, which appear to be simple, are really complex and require judgment for interpretation. It is the purpose of this chapter to consider the various details of the mechanical make-up of tests and to show their implications.

The identification of the pupil taking the test. So far as the interpretation of the results of any test is concerned, the pupil is just as important as the test that he takes, and his identification is therefore of real importance. The first page of a test (or some portion of the first page) usually contains somewhere on it spaces for the necessary data concerning the identification of the pupil. All the spaces should be filled in, or if the information desired is not available a dash should be made to indicate that the information desired has not been neglected. It is doubly important that all the data should be supplied if possible when the tests are to be scored or otherwise used by persons other than the classroom teacher who is familiar with the pupils.

1. *The name of the pupil.* This is an important means of identification and is usually stressed on the test sheets. Practices differ as to the writing in of the pupil's name. Two points should be checked by the teacher in charge: (1) that the name is written legibly and (2) that the surname is clearly indicated. If the papers are to be used by another scorer it is a good practice for the teacher to examine each paper when it is completed and to underline in blue or red pencil the last name of each child. If the names were to be written by the children, some would write their surnames first and others their given names, and frequently they are difficult to distinguish. In the lower grades it is not unusual for children to write in only their first names. Of course this means that the teacher should supply the surnames when checking the papers, so that a reader unfamiliar with the pupils will have sufficient information.

2. *The sex of the pupil.* Under ordinary conditions this is a relatively unimportant item unless the pupil's sex is to be used as a basis for some use of the test or of results of the tests. This might happen in the grammar grades or in the junior high school if children were to be divided on the basis of sex in a reclassification. Most children know their sex, although occasionally a pupil in the first or second grade does not, and the names are not always self-explanatory. Here again is a possible deficiency which the teacher should complete.

3. *The age of the pupil.* The calculation of the age of pupils is one of the real difficulties in the accumulation of the data about children, since common phraseology is so variously understood. To say that a pupil is ten years old may mean that he is exactly ten years old or that he is anywhere from nine years and six months to ten years and six months old. It usually means that he is over ten years and just less than eleven years.

Chronological age is one of the more important factors in determining the judgments concerning a pupil, and as such it should be carefully determined and be as free from error as possible. As a rule ages are kept by school-administrators in units of six months or more. The technics for the use of test results, however, require a knowledge of ages much more exact than that. To determine these exact ages it is necessary to know more than the answer to the question "How old are you?" The present practice is to ask two questions, since most children do not know the exact date of their birth. The first question, one which most children can answer, is "What was your age on your last birthday?" They can then say that they were ten years old, eleven years old, and so on. The second question, "When is your birthday?" should bring the answer "April 15," or "February 19," and so on. From this data exact ages for testing purposes can be calculated.

There are various ways of calculating the age of children in months. The most logical, of course, is to calculate the age from the birthday to the date of the test and take the nearest whole number of months. By this means of calculation the chronological age of pupils would vary from fifteen days before the date of the test to fifteen days afterwards.

Let it be supposed that a child was ten years old at his last birthday, that his birthday was March 21, and that a test was given on May 10. Calculating to the nearest month, there would be ten years plus two months, since the month would fall between April 26, fifteen days before the test, and May 25, fifteen days after the test. The two-month period would be from March 21 to May 21.

This method of calculating is probably more frequently used than other methods, but when there are large numbers of pupils it is more laborious than to calculate to the last completed month of age. In the latter case it is merely

necessary to compare the date of the birthday and the date of the test without taking into consideration a period before and after the test.

In the case just cited, a child ten years old on March 21 would on May 10 have completed only one complete month beyond his birthday (from March 21 to April 21), and by the second method he would be reckoned as ten years and one month old. The latter method, although it decreases the chronological age of practically half the pupils by about one month over the former method, is much more simple to work out and is likely to give as good comparative results.

When the teacher is concerned only with his own pupils, either method of calculation is justifiable. When the results of an entire school system are to be combined, it is wise for the chief examiner to define his method of reaching the ages of pupils and to see to it that all the ages are calculated in the same way.

4. *The school grade of the pupil.* As a rule there is little difficulty in the correct placement of this information, although there are instances where a strange examiner may find difficulty in interpreting the meaning of what may be given. Where there are several classes of the same grade in the same building it is not unusual to find them distinguished by letters; as, IV-A, IV-B, IV-C, and so on. Sometimes the letters have no meaning other than that they are separate classes of the same grade, and, if questioned, the teachers would say that the two or three classes are equivalent. Again, it is sometimes the practice of schools where these letter ratings are used to divide the children of a grade into classes in accordance with their supposed abilities, indicating the advanced group as A, the poorest group as C, and the middle group as B. Yet again, when semiannual promotion is possible, a class called IV-A may mean a group beginning the first half of fourth-grade work, and IV-B may indicate

a class beginning the second half of fourth-grade work. When letters are used on the test papers to denote grades it is a good practice for the classroom teacher to indicate on some sheet which is included with the package of test papers sent to the scorer just what is meant by these letters.

5. *The school of the pupil.* This is an important piece of information, particularly if the results of tests in a number of schools are to be tabulated. It is not enough to place this information on one paper only, since the test papers are frequently separated in any large survey. When the papers of any class are to leave the school building in which they are given, it is always wise to write the name of the school on each test or test booklet.

6. *The name of the teacher.* This information is also of value and becomes of special importance when (as is frequently the case with intelligence tests) the results are to be carried on for use in later years. It is also of importance when there are in the same school building several classes of the same grade, and where these classes are not given separate designations, such as letters. Then the name of the teacher is the quickest means of locating the placement of any paper.

Fig. 29 shows the information for the identification of the pupil and school which is called for on the National Intelligence Test.

The identification of the test. The identification of the test itself is an element which each teacher who uses it should note and check. It is an easy thing to do, but the failure to do it may lead to difficulties in the scoring or interpretation of the data.

1. *The subject of the test.* This data is usually placed in a prominent position on the first page of the test. It is frequently combined with the name of the author or of the group or institution which publishes or sponsors it. Thus

NATIONAL INTELLIGENCE TESTS

SCALE A — FORM 3

Prepared under the auspices of the National Research Council by M. E. Haggerty, L. M. Terman, E. L. Thorndike, G. M. Whipple, and R. M. Yerkes

Name _____ *Grade* _____ *Boy or girl* _____
First name Last name

Date of birth _____ *Age* _____
Month Day Year Years Months

Birthplace of parents _____ *Race* _____
Father Mother

Name of teacher _____

Name of school _____

Name of city _____

Date of examination _____

FIG. 29. Form for the identification of the pupil from the National Intelligence Test, Scale A, Form 3[1]

there are the Woody Arithmetic Test, the Morrison-McCall Spelling Scale, the Cleveland Survey, and the like.

2. *The forms or parts.* Tests are frequently so arranged that there are several tests of equivalent or nearly equivalent difficulty from which a teacher may choose; and frequently there are several different tests which form a series, and there may be alternative forms of each type of test within the series.

In the Spencer Diagnostic Tests in Arithmetic there are three different tests which form a series. Test I is of one type, Test II of another, and Test III of still another. Each of these types of diagnostic tests is published in two forms, A and B. The Morrison-McCall Spelling Scales are eight

[1] Reproduced by permission of the World Book Company.

in number and are of equal difficulty, so that they can, if necessary, be used in rotation with the same class, without repetition.

The intelligence tests are published in several different scales and forms as well. The National Intelligence Test, for example, is published in three different forms for each of two different scales, so that six tests are available for testing the same set of abilities.

The fact that tests are now appearing in interchangeable forms makes it possible for a teacher to retest pupils at such intervals as may be thought wise without any danger that the knowledge received from one test will carry over to another. A child may improve in the technic of taking tests through practice in taking them, but as a rule the technics are so simple that such practice is of relatively small advantage. The fact that there are several forms makes it necessary for the examiner to check carefully on the identification of the test about to be used, to make sure the test has not been already given to the same children, and that the form in hand is the one which it is desirable to use.

Recording and totaling the final scores. On the first sheet of a test which comprises several pages, there are usually spaces for recording and totaling the final scores. In some types of tests (for example, the intelligence tests) this is a somewhat complicated process, since the final scores are derived from treating the scores of the different parts in different ways. On other tests it is merely the total scores of the separate parts which, when added together, give the final score.

In Fig. 30 is shown the space for recording, adjusting, and totaling the scores on the National Intelligence Test. It should be noted that the several scores as they are transferred from the separate parts of the test are treated differently, and are added only for the final scores when they are adjusted. The numbers in the spaces are a pupil's scores.

Directions for giving and taking the test. There are usually two sets of directions supplied with tests, although many tests include all the necessary directions on the examination sheets. In any case there are the directions to guide the examiner in giving the test and the specific directions for the pupil in taking it. In the non-language tests all the directions for taking the test are intended for dictation by

Test	*Rights*	*Method*	*Score*
1	9	$\times 2 =$	18
2	15	$\times 2 =$	30
3	31	$=$	31
4	23	$-\frac{\text{Wrongs}}{(14)} =$	9
5	67	$\times {}^{3}/_{10} =$	20.1
Total		*Sum*	108.1

FIG. 30. Form for recording scores (with pupil's answers) on the National Intelligence Test, Scale A, Form 3 [1]

the examiner or the teacher; in the language forms there are separate directions for the examiner and the pupils.

The directions to the examiner specify how the test shall be given to the pupils, the amount of time that should be allowed for it, any special directions that should be given, any special preparations that should be made, and, usually, the exact wording that the examiner should use in introducing and administering the test or parts of it. Such a series of directions is given in the directions for using the Thorndike-McCall Reading Scale, as shown in Fig. 31. They are not intended to be read to the pupils.

[1] Reproduced by permission of the World Book Company.

The directions which are printed for pupils to follow are usually simple and clear. They rarely need amplification on the part of the examiner, and should have none unless the directions to the teacher specifically say that such amplification may be given if necessary. Samples of directions are "Read this and then write the answers," "Read again if you need to," "Read each question and draw a line under the right answer," "In each sentence draw a line under the

HOW TO APPLY TEST

To the Examiner: Distribute test booklets with front page up. Request pupils not to open booklets until the signal is given. Have pupils fill out the blanks at top of the front page. Read the paragraph aloud while pupils read silently. Read the first question aloud. Have it answered orally and then in writing by the pupils. Stop pupils thirty minutes after saying *Open paper! Begin!* Give no further help.

FIG. 31. Directions to examiners from the Thorndike-McCall Reading Test[1]

one word that makes the sentence true," "Work as many of the examples as you can before you are told to stop."

When all the directions are for the examiner, and the directions for the pupils are to be dictated by him, a booklet of directions to be followed by the examiner usually accompanies the test. This booklet will be referred to in a later section, since it contains other materials as well. Such a series of directions is found in the Pintner-Cunningham Primary Mental Test, which consists of a number of pages of pictures made in booklet form. The directions for the early portions of the test are reproduced in Fig. 32. The words in boldface type are to be read to the pupils; the other materials are for the guidance of the examiner.

[1] Reproduced by permission of Teachers College Bureau of Publications.

DIRECTIONS FOR ADMINISTERING

The number of kindergarten children that should be tested at any one time will depend upon the skill of the examiner and the amount of assistance available. If the examiner is alone, probably not more than ten or fifteen children should be examined together. If there are one or two assistants to help in seeing that the children have turned to the right page, then a group of twenty or thirty may be easily handled. In the first and second grades there is no problem as to the number who can take the test at any one time.

Seat children so as to minimize chances for copying from one another. Provide the children with soft pencils or sharp crayons. See that each child's name is written on his booklet. Additional data as to age, grade, etc., should be obtained from the teacher at another hour. Pass booklets to children, asking them to keep the books closed until told to open them. Be very deliberate in giving all directions. Everything that is to be said to the children is printed in boldface type. Do not read anything else to the children.

When books have been distributed, say: **We are going to play some games with the pictures in this book. We are going to mark in this book. Listen carefully to everything that I say so that you will mark just the right things. Open your books** (illustrate by turning to page 2) **and fold them back like this, so that you can see an apron.** (See that all have the right page.)

Test 1. Common Observation (pages 2 and 3)

Page 2. **Look at the pictures.** (Indicate the first row, page 2.) **We are going to put some marks on some of the pictures, but not on all of them. Let's mark the things that Mother uses when she sews her apron. Can you find the apron? Put your finger on it. Mark it like this.** (Draw line on board.) **Mark just the apron and not anything else.** (See that all have followed directions. Any kind of mark that indicates the object is satisfactory.)

FIG. 32. Directions for giving the Pintner-Cunningham Mental Test[1]

[1] Reproduced by permission of the World Book Company.

The preliminary practice test. In many tests as they are now published, opportunity is given for the pupils who take the test to have some practice in the particular technics before they actually attempt the test proper. This sort of practice is unnecessary in certain forms of tests, such as in spelling and arithmetic, although it is occasionally used in these. The practice element is valuable; for children are found who have not taken similar tests and do not know the technics, and others are found in the same classes who have had the opportunity to learn them. A small amount of practice enables all the pupils to have an equal opportunity to do the required work without handicaps. When the child misinterprets the directions for doing the test, and there is no practice test which would enable him to correct his difficulty, he is frequently at a greater disadvantage than either his ability or his achievement would warrant. It is true, of course, that the directions in most tests are in reality a part of the tests, but it is also true that the makers of tests are placing more and more emphasis on the results to be obtained where the technics of taking the tests are equalized for all pupils so that differences in achievement will be clearly shown.

The practice element in the Thorndike-McCall Reading Scale, which was referred to in a preceding section, is shown in Fig. 33. It provides a practice form similar to the actual forms which the pupils are later expected to complete.

The practice element in one of the intelligence tests is given in Fig. 34 on page 105.

The preliminary test, or practice test, is given for the purpose of making sure that the pupils understand the technic of what they are to do, when that technic contributes to the amount of product which the pupil can make, but when it is not also a part of the test itself. When the technic is in reality a part of the test itself, there can be no preliminary test, and the examiner must not allow any.

In handwriting it is the purpose of the examiner to try to find a fairly typical sample of the pupil's handwriting. If this sample is taken from a paragraph written from dictation, the handwriting will not be alike throughout the paragraph. The first few words or sentences may be written with

This is to be a reading contest. You will read paragraphs like this one, and answer questions like those you see below. Answer every question you can. If you come to a question you can't answer, skip it and go on. Go back to it later. If you finish before you are told to stop, go back and make sure you have made no mistakes. When possible, the answers to the questions must be found in the paragraph. You may read the paragraph as many times as you need to. You will have enough time, but don't waste it. Play fair. Don't look at anyone else's paper. You will be told your score later.

I. Does it say you are to do your best?_____

II. Does it say you may read the paragraph as often as you need to?_____

III. Are you to look at anyone else's paper?_____

IV. When possible, where must you find the answers?_____

_____

As soon as you finish one page do the next. You are to start with the first page. Open paper! Begin!

FIG. 33. Practice element, Thorndike-McCall Reading Scale[1]

meticulous care, and the last sentences may be written hurriedly and carelessly. Obviously neither of these parts is typical of the pupil's handwriting, and therefore the middle of the paragraph is selected for comparison with the scale. The first part of the paragraph written bears some analogy to the practice test such as has been described.[2]

[1] Reproduced by permission of Teachers College Bureau of Publications.

[2] See page 103.

The body, or subject matter, of the test. This part of the standard test is the part most familiar to the average person, and from many points of view is the center of interest. Samples of this portion of the tests have already been given. They vary in character, form, and make-up, but are designed

Exercise 1

Find all the answers as quickly as you can.
Write the answers on the dotted lines.
Use the sides or bottom of the page to figure on.

Begin here

1 How many cents are six cents and five cents? *Answer*____

2 A girl earned 75 cents and spent 43 cents. How much did she have left? *Answer*____

3 How many nickels make a dollar? *Answer*____

4 How many square inches are there in a card 7 inches long by 6 inches wide? *Answer*____

5 How long will it take a man to walk 19 miles at the rate of 4 miles an hour while walking if he makes two stops of an average length of 15 minutes each? *Answer*____

6 On December 20 the sun rises at 7:22 and sets at 4:48. How much less of the day is daylight than dark? *Answer*____

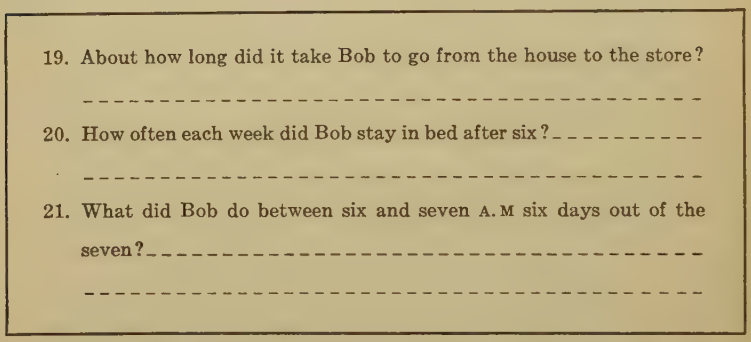
FIG. 34. Practice element, Exercise 1, National Intelligence Test, Scale A, Form 3[1]

to carry out the purposes for which the tests are proposed, and each element in them is the result of careful selection, scrutiny, and evaluation.

Provision for the answers to the test elements. It has been stated that one of the greatest advantages the standard tests have over other forms is the greater degree to which they are objective. This greater objectivity simply means that different examiners are more likely to give the same judgments

[1] Reproduced by permission of the World Book Company.

concerning the same answers, or that the same examiner is more likely to do so. This objectiveness is secured by many different devices, but almost always through a device which narrows the field of the answer to short, definite, and easily interpreted reactions. In certain forms of arithmetic tests the reactions are the time-honored answers to the problems; in reading tests the reactions are in terms of some action on the part of the pupil or upon some written word or words to


19. About how long did it take Bob to go from the house to the store?

_____

20. How often each week did Bob stay in bed after six? _____

_____

21. What did Bob do between six and seven A.M six days out of the seven? _____

_____

FIG. 35. Answer spaces from the Thorndike-McCall Reading Scale, Form 2[1]

indicate that reaction; in other tests the reaction may be indicated by the underlining of a word, by filling in a blank space, by checking an answer from among several that are given, or by crossing out some element; and in such tests as those in spelling the answer is indicated by writing the word dictated by the examiner.

When the answers are short and pithy, provision can be made for them by leaving spaces in which to write them. This is the case in the Thorndike-McCall Reading Scale, part of which is reproduced in Fig. 35.

When there are only two possibilities in the answer, such

[1] Reproduced by permission of Teachers College Bureau of Publications.

as "Yes" or "No," the words may be written in, and the technic for recording the reactions may be to underline the correct answer; when there are two or more possible answers, they may all be included on the test sheet, and the technic may be to underline the correct answer. Incorrectness is usually not emphasized.

In primary tests, where handwriting on the part of the pupils is not easily done, the reactions are sometimes reached by drawing lines, by making additions to pictures, and the like. The reactions in the Haggerty Reading Test, Sigma 1, for Primary Grades are shown in Fig. 36.

Spaces for totaling the scores of the test units. Somewhere on each page of a test there should be found a space or blank form to be used for the insertion of the test score received by the pupil for the achievement made on that page. Where such a formal space is not provided, there is usually a free margin or other space available which it is expected will be used for that purpose.

Directions for the examiner to use in the correction or scoring of tests. A test manual, or manual of directions for the use of the examiner, is as a rule included as a part of every test sent out by the publisher or distributor with any order for tests. Frequently, as has been stated, this manual contains general directions for giving the test; it also should contain definite directions for correcting and scoring it.

The method of correcting tests may vary with each test. In some cases where there is a wide variety in the answers considerable emphasis must be laid on the judgment of the scorer as to the exact interpretation of any answer, and therefore detailed directions must be included. Fortunately this type of answer is gradually being eliminated from the results of testing as the form of questions is becoming more definite, and the technic of making tests helps to make the elements more objective.

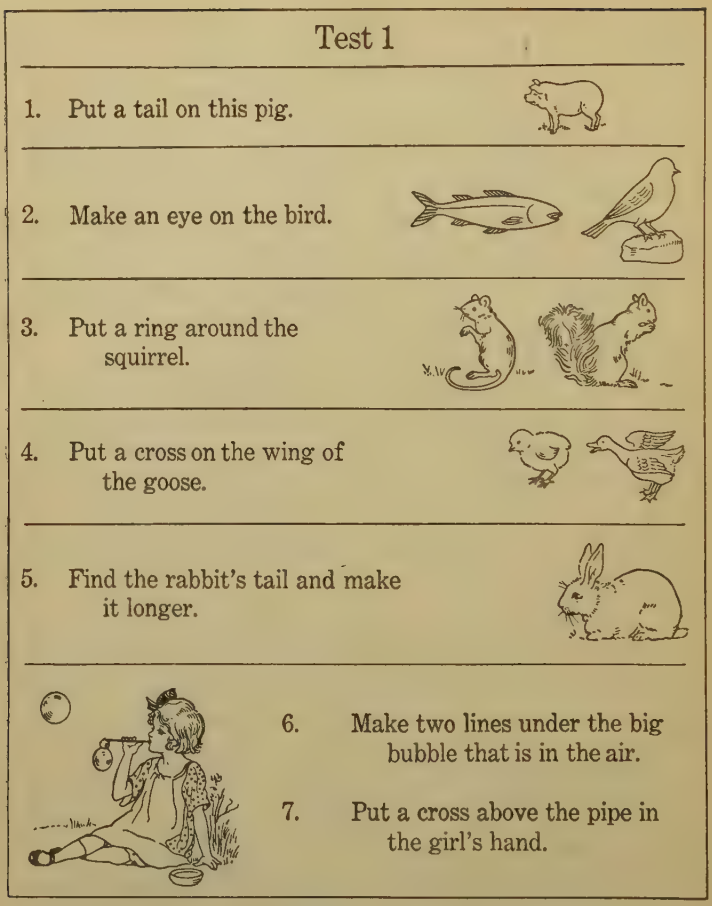

Test 1
1. Put a tail on this pig.
2. Make an eye on the bird.
3. Put a ring around the squirrel.
4. Put a cross on the wing of the goose.
5. Find the rabbit's tail and make it longer.
6. Make two lines under the big bubble that is in the air.
7. Put a cross above the pipe in the girl's hand.

FIG. 36. Portion of Haggerty Reading Test, Sigma 1[1]

Where the interpretation or judgment of the teacher is involved in scoring an answer, the test manual will list all the correct answers. It will also frequently list a variety of wrong answers to each item, so that the scorer can judge

[1] Reproduced by permission of World Book Company.

SCORING KEY

The only portion of the answer required for correctness is in italics.

*Samples of incorrect answers.

1. The salt was in a *bag*.
2. She bought a *dozen* eggs; *12*.
3. The children seen a *fire-engine* with three horses.
4. *Water* Street; *Walter* Street.
5. He *killed* the *fox*; *Kill it*.
   *Killed; *Shot him.
6. The rabbit was *brown*.
7. *In* the *woods*; *In* the *forest*.
   *Woods.
8. Beauty had *three* brothers; *Three* boys; *3*.
9. *Yes*; *Younger* than both; She was *younger*.
10. *No*; *No, Beauty*; The *youngest*.
    *Beauty; *Young.
11. *Jealousy*; *Jealous*.
12. *Music* lessons; *Piano lessons*; *Play piano*.
    *Grace's music; *Singing; *Piano; *Music and hens.
13. *No*; Grace *sell* them; *Not*.
14. *Yes*; *yes*; *He doesn't like it at all*.
    *Fred does not like music.
15. *Richard*; *Richard* and Edward
    *Edward; *Richard and Henry.
16. *Sam* and *Henry*; *Sam* and *Edward*.
    *Edward doesn't like Sam; *Sam, Henry, Edward.
17. *Arthur* and *Richard*.
18. *Henry* and *Edward*; *Henry*; *Edward*.
19. *One Hour*; From *ten minutes of seven to ten minutes of eight*; He *left* the *house at ten minutes of seven* and *reached* the *store at ten minutes of eight*.
    *Almost an hour.
20. *Once a week*; *One time*; *One day*; *Every Sunday*; *Only Sunday*; *On Sunday*; *Sundays*; *At Sunday*; *1 each week*; *1 a week*.
    *Sunday; *One.

FIG. 37. Portion of scoring key from the Thorndike-McCall Reading Scale, Form 2[1]

[1] Reproduced by permission of Teachers College Bureau of Publications.

more easily. An instance of this is found in the scoring sheets for the Thorndike-McCall Reading Scale, where correct answers are printed in one style of type and wrong answers in another, as in Fig. 37.

In other types of tests it is sometimes possible for the answers to be listed in a scoring key. The scoring key is a list of the correct answers for a test placed in such form that it can be either superimposed on the answer sheet of the pupil, or placed beside it and direct comparison made of the pupil's answers with those on the key. A scoring key used by superposition for the National Intelligence Test is printed on transparent paper. It places the answers in their proper positions on the sheet and makes it possible to compare every element of the pupil's sheet with the sheet of the key. Such a transparent sheet is illustrated in Fig. 38 for Test 5 of the National Intelligence Test, Scale A, Form 3.

Another type of key, given in many tests, can be laid beside a column of answers, and a check enables the scorer to determine which of the pupil's answers are like the key and which are different. Such a key is shown in Fig. 39.

Norms for the test. Norms are the levels of achievement that represent the result of determining what pupils have done in a given situation, at a given age, in a given community, in a given section of the country, or in the country as a whole. When tests are given to a large number of pupils the results are tabulated. From this tabulation the average achievements of different types of pupils can be determined, and these ratings are then published as guides for the examiners in judging the work of any pupil or group of pupils.

Norms are derived in various ways, and the lists of norms that are published in test manuals or in direction sheets usually tell how they are derived. Norms may be derived from a certain type of school in a certain type of community (thus, norms published may be derived from scores made

KEY FOR SCALE A — FORM 3

Second Sheet

TEST	1	2	3	4	5	TOTAL
MAXIMUM SCORE	32	40	48	40	36	196

Test 5

				(5)				(10)					(15)				(20)
3	1	2	1	2	3	2	4	3	5	3	7	1	2	6	8	9	5	4	2
				(25)				(30)					(35)				(40)
5	1	9	7	8	3	4	6	3	5	7	9	8	4	2	6	1	2	6	1
				(45)				(50)					(55)				(60)
9	4	5	8	7	3	5	6	8	1	7	9	3	2	4	7	4	3	5	8
				(65)				(70)					(75)				(80)
6	1	9	2	6	2	1	8	3	5	9	4	7	3	9	1	7	4	2	5
				(85)				(90)					(95)				(100)
6	8	1	5	7	9	3	2	1	8	4	6	9	7	4	6	3	5	8	7
				(105)				(110)					(115)				(120)
9	6	4	8	5	1	9	6	8	7	1	2	4	9	7	5	6	8	3	5

FIG. 38. Key sheet from National Intelligence Test, Scale A, Form 3, Test 5[1]

[1] Reproduced by permission of the World Book Company.

by pupils in cities of over twenty thousand population); they may be derived from scores made by pupils in certain sections of the country (for example, by pupils in the North Atlantic states), or they may be derived from all the scores made by all pupils without regard to where the tests have been given.

OTIS CLASSIFICATION TEST
KEY FOR FORM A, PART I

1. (2)	56. (5)
2. (3)	55. (30)
	54. (4)
3. (5)	
	53. (5)
4. (4)	
5. (2)	52. (2)
6. (3)	
	51. (9)
7. (3)	
8. (1)	50. (7)
	49. (54)
9. (2)	48. (16)
	47. (13)
10. (2)	46. (3)
11. (250)	45. (1)
12. (200)	
	44. (3)

FIG. 39. Scoring key, Otis Classification Test, Form A, portion of Part I[1]

Norms are also given in terms of different units. Some are listed in terms of age. Age norms are the average attainments of pupils of different given ages. Thus when tests have been given to pupils in many different communities the scores of twelve-year-old pupils, regardless of the grade they may be in, are assembled; so also are the scores of all pupils eleven years old, ten years old, and so on, and from these scores which pupils have actually made can be determined the average achievement of children of each of these ages. Such a list of age norms is given in Fig. 40.

Norms are also given in terms of grade. Just as age norms are derived from the actual scores of all pupils of a given age, so grade norms are derived from all the scores of pupils in any given grade. The scores of fourth-grade pupils, of

[1] Reproduced by permission of the World Book Company.

AGE NORMS IN ACHIEVEMENT TEST, MENTAL ABILITY TEST, AND CLASSIFICATION TEST (TOTAL SCORE)

Years:	8			9			10			11			12			13			14			15			16			17			18 or over		
MONTHS	ACH.	M. A.	CLASS.	ACH.	M. A.	CLASS.	ACH.	M. A.	CLASS.	ACH.	M. A.	CLASS.	ACH.	M. A.	CLASS.	ACH.	M. A.	CLASS.	ACH.	M. A.	CLASS.	ACH.	M. A.	CLASS.	ACH.	M. A.	CLASS.	ACH.	M. A.	CLASS.	ACH.	M. A.	CLASS.
0	6	8	13	16	15	31	26	23	49	36	31	67	45	38	83	53	44	97	59	49	108	63	53	116	67	56	123	70	58	128	71	59	130
1	7	8	15	17	16	33	27	24	51	37	32	69	46	39	85	53	44	98	59	49	109	64	53	117	67	56	124	70	58	128	.....		
2	8	8	16	18	16	34	28	24	52	38	32	70	47	39	86	54	45	99	60	50	110	64	53	117	68	56	124	70	58	129	.....		
3	9	9	18	19	17	36	29	25	54	38	33	71	47	40	87	54	45	100	60	50	110	64	54	118	68	57	125	71	58	129	.....		
4	9	10	19	19	18	37	29	26	55	39	34	73	48	40	88	55	46	101	61	50	111	65	54	119	68	57	125	71	58	129	.....		
5	10	10	21	20	18	39	30	26	57	40	34	74	49	41	90	55	46	101	61	51	112	65	54	119	68	57	125	71	58	129	.....		
6	11	11	22	21	19	40	31	27	58	41	35	75	49	41	91	56	46	102	61	51	112	65	55	120	69	57	126	71	59	130	.....		
7	12	12	24	22	20	42	32	28	60	42	35	77	50	42	92	56	47	103	62	51	113	66	55	121	69	57	126	71	59	130	.....		
8	13	12	25	23	20	43	33	28	61	42	36	78	50	42	93	57	47	104	62	52	114	66	55	121	69	58	127	71	59	130	.....		
9	14	13	27	24	21	45	33	29	63	43	36	79	51	43	94	57	48	105	62	52	114	66	55	122	69	58	127	71	59	130	.....		
10	14	14	28	24	22	46	34	30	64	44	37	81	52	43	95	58	48	106	63	52	115	66	56	122	70	58	127	71	59	130	.....		
11	15	14	30	25	22	48	35	30	65	45	37	82	52	43	96	58	49	107	63	53	116	67	56	123	70	58	128	71	59	130	.....		

FIG. 40. Age norms, Otis Classification Test, Form A[1]

[1] Reproduced by permission of the World Book Company.

fifth-grade pupils, of sixth-grade pupils, and so on are assembled, and from them can be found the average achievement of pupils in any given grade taking any given test. When grade norms are published it is stated at what time of the year the tests were given and the norms derived. This is important information and, as will be shown in a later chapter, is necessary, in order to use the norms for comparing the work of pupils in one place with the work of those in other places. Fig. 41 gives a list of grade norms for a standard test.

Woody-McCall Standards for the Beginning of the School Year for Forms I, II, III, and IV [1]

For the grade as a whole				For high and low sections separately						
Grade	Standard	Number of Cases		Grade	Standard	Number of Cases	Grade	Standard	Number of Cases
III	6.8	6,665		III low	6.1	1,244	VI low	22.0	2,601
IV	13.1	12,572		III high	10.6	942	VI high	24.3	2,240
V	17.8	11,416		IV low	12.5	2,635	VII low	25.4	2,574
VI	22.5	10,488		IV high	16.4	2,322	VII high	27.4	1,847
VII	25.9	9,074		V low	17.2	2,395	VIII low	27.9	2,060
VIII	27.8	6,174		V high	19.9	2,150	VIII high	28.5	1,328

To make the Woody-McCall standards comparable with results obtained later in the school year for each month after October, add the following increments

III	.54	V	.42	VII	.25
IV	.43	VI	.24	VIII	.20

[1] Woody's revised standards of March, 1923. Two sets of standards are given: For each grade as a whole, and for each half-grade. Both sets of standards are based upon the weighted averages of scores for the different grades when reduced to an October basis. The apparent incongruities in the two sets of standards are due to the fact that different groups are considered in forming the two sets of standards. The system of increments for the purpose of making the standards comparable with results obtained later in the school year involves only a small amount of error, and may be applied to both sets of standards.

FIG. 41. Grade norms, the Woody-McCall Arithmetic Test, Mixed Fundamentals [1]

Norms are constantly being revised for all tests as more and more data become available; therefore it is always wise to make sure that the norms to be used are the most recent published and also that they are based on as many cases as possible. The greater the number of cases the more likelihood

[1] Reproduced by permission of Teachers College Bureau of Publications.

there is that the norms are correct, and the more useful they become in interpretation of results.

Directions for preparing the scores for use and interpretation. With many tests there is published as a part of the test manual a series of directions for the use and interpretation of the scores. In diagnostic tests these directions show how to use the scores obtained for the location of difficulties and errors; they also show the significance of various types of achievement. In intelligence or ability tests these directions show how to transpose the scores into other and more useful scores, such as the mental ages or intelligence quotients. A study of this portion of the manual for any test will aid a teacher materially in finding what the test has measured and will offer suggestions for uses of the scores which may not have been anticipated.

Scoring and graph sheets. With test manuals or with groups of tests it is sometimes the practice of publishers to include a special form of graph sheet or a scoring blank upon which to record the results of the testing. This sheet is made up especially for the test which is in use and is constructed in a form that makes the tabulation of data easy for the teacher. In some tests this sheet shows the norms in graphic form; in others the norms for the data must be taken from a table. The purpose of these sheets is simply to enable the teacher or examiner to assemble his data more easily and quickly than otherwise would be possible, and is usually of great help in that it makes for economy of time and effort, and at the same time promotes accuracy in the assembly and completeness of data.[1]

Report to the author. Norms and standards are made up, as has been shown, from the actual results of the tests in various communities or sections of the country. These results must be obtained from the users of tests, and are

[1] A sheet of this type is shown in Fig. 98, p. 379.

modified in terms of the results which are received. For this reason it is the practice of many authors and publishers of tests to include among the test materials a sheet upon which the results of testing in any community or class can be assembled as a report to the author. This data, when used by the author in assembling new norms for the tests, aids in making the results of testing more useful for others.

Chapter summary. When tests are distributed they contain many materials and include much more data than the subject matter of the tests themselves.

One of the more important elements of the tests is the provision which is made for the identification of the pupil taking the test. This identification consists of name, sex, age in months, school grade and section, school, and teacher. All these data are necessary for complete identification, and should be checked by the examiner either at the time of giving the test, or later from the teacher's class book or records.

A second element in the construction of a test is the identification of the test itself. Since the same test may be published in several different forms so that its values may be realized more than once with the same group of pupils, it is always necessary for the teacher to know what tests and what forms or parts of tests have been previously used with the pupils.

A third element in the construction of tests consists in the spaces included for the recording and totaling of the final scores on the tests. When there are no such spaces, it is expected that margins or other free space on the test sheets will be so used.

A fourth element consists in the directions for giving and taking the test. As a rule the directions for giving the test, which are for the use of the examiner, are given in the test manual or booklet of directions that is supplied with the test materials. Generally the directions for taking the test, which

are for the use of the pupils, will be found on the test sheets preceding the subject matter of the test. In tests that are to be dictated, such as spelling tests, or where the directions are to be dictated, as in primary mental tests of various kinds, the directions for taking the test are incorporated in the test manual as a part of the directions for the examiner.

In many tests a fifth element is included, the preliminary practice test, which is a device utilized when it is desired to eliminate the technic of taking the test from the achievement of the pupil. This preliminary practice of the test elements may appear in several forms. It is usually a miniature of the test itself, and contains more or less obvious material. It is rarely counted in the final test score.

The subject matter of the test itself forms a sixth element in the construction of tests and is that part of the test with which most teachers are familiar.

An important item in the construction of a test is the provision which is made for the answers to the test elements. The more definite and simple these provisions are, the more likelihood there is that the test will be objective in its scoring and accurate in its results.

An eighth element is the space which is allowed for the placement of the scores on the various test units. This is somewhat similar to the formal space for the totaling of the final scores; and when it is not included, it is expected that otherwise unused space in the margins or elsewhere will be found for the purpose.

A most important part of the test material is that devoted to the directions for the examiner to use in the correction and scoring of the tests. These directions are usually found in the test manual which accompanies the test materials and contains samples of the right answers, either in the form of lists or of keys, and, also, where there is a possibility of variability in answers, a list of various forms of wrong answers.

Another element of great value to the examiner and teacher are the norms for the test, the use of which will be described in later chapters. Norms are derived from actual results of testing, with certain types of schools, from certain types of communities, from given sections of the country, or from the results of all the tests regardless of the source. Norms, or standards, may be in terms of age or grade, or both. They simply indicate the average achievement indicated by the results of the tests.

Another element within the test materials consists of the directions for preparing the scores for use and interpretation. Some tests need special attention or emphasis on certain elements for the best interpretation of the scores, and directions are provided for that purpose.

Scoring and graph sheets frequently accompany the other test materials and are included to aid the teacher in assembling the data more easily and quickly.

A final element in the construction of tests which is included with many tests is a sheet to be filled out by the examiner at the conclusion of the testing, as a report to the author of the scores or achievement of the pupils. This report is used by the author in determining, correcting, adding to, or improving the norms and standards for future publication.

SELECTED BIBLIOGRAPHY

McCall, W. A. How to Measure in Education, Part Two. The Macmillan Company, New York, 1922.

Monroe, W. S. An Introduction to the Theory of Educational Measurements. Houghton Mifflin Company, Boston, 1923.

Monroe, W. S., DeVoss, J. C., and Kelly, F. J. Educational Tests and Measurements (Revised and Enlarged Edition), chap. xi. Houghton Mifflin Company, Boston, 1924.

Otis, A. S. Statistical Method in Educational Measurement, pp. 160–174. World Book Company, Yonkers-on-Hudson, 1925.

CHAPTER VI

THE GIVING AND SCORING OF TESTS

Ways of giving standard tests. There are several ways in which tests may be given, varying with the types of the tests, the forms in which they are published, and the uses to which they are put.

One way of giving tests is by dictation of the entire test to the group of pupils. This is used with spelling tests or when products are desired for comparison with a scale, as in handwriting and composition. The directions used in giving spelling tests differ somewhat according to the particular test in use, but in general spelling tests consist of the dictation of a list of words to the pupils. The pupils are asked to prepare a sheet of paper by numbering it from top to bottom along the left-hand margin, one number to a line, beginning at a certain line at the top, and the examiner then dictates the words of the test, allowing a suitable interval between words. One of the difficulties encountered in this method of giving a spelling test is that the pupils may make errors in numbering the lines and place the words out of their proper order. A teacher should therefore see to it that the preliminary directions are followed by the pupils in each step. A second difficulty arises because of the fact that (in the lower grades particularly) a portion of the test is considered by the teacher to be too difficult. It is undoubtedly too difficult for the majority of the pupils; but, regardless of the grade or the feelings of either the teacher or the pupils, the entire test should be dictated. A majority of the pupils in the third grade may stop work completely when the test is half dic-

tated, and just sit quietly during the remainder of the test, but some pupils will attempt the words, and will frequently be able to spell correctly words which the teacher would not have expected that they could spell. In a test recently given to a group of third-grade pupils the word "superintendent" appeared far down on the list, and several pupils spelled it correctly. If the test had been stopped when the majority of the pupils had given up, the few who could spell "superintendent" would not have had the opportunity to prove their knowledge.

When dictating a test, particularly a spelling test, the examiner should be careful to enunciate clearly, to speak slowly, and to keep his face turned toward the class when he is speaking to them. One of the greatest faults with examiners who are dictating either directions or tests is that they have a tendency to keep the head bent down when they are reading. The teacher should be familiar enough with the materials to speak to the class so that every child is enabled to hear clearly. This is particularly necessary when the examiner is not familiar to the class.

A second way of giving tests is through the medium of printed booklets or sheets to be filled out by the pupils and passed to the examiner at the conclusion of the testing period. The method of giving the test in this case is to distribute the booklets or sheets so that none of the test material can be seen before the opening of the test period and then, by a series of dictated directions, to tell the pupils what they are expected to do and when the allotted time has been completed. In many of the printed tests there are few directions for the examiner, and his work during the progress of the test is simply to see that the test is going properly. In other forms of these tests the examiner must not only start the test but must also add directions from time to time in accordance with the directions given in the test manual.

In the Thorndike-McCall Reading Scale the examiner presents the materials of the first page, which are in the form of a practice test. When this has been completed the pupils are directed to turn over the first leaf and to continue to work until told to stop. An even simpler procedure is used in the Woody-McCall Arithmetic Test, where the directions are printed, and all that the examiner needs to do is to have the pupils turn the sheets to the printed side and go to work.

The tests which take the greatest care on the part of the examiner are the group intelligence tests, where each portion of the test is timed and where each portion must have a different type of direction given. In the National Intelligence Test the examiner has the pupils turn to a page headed "Exercise 1." The pupils practice this exercise for one minute, and then the examiner directs them to turn the page and continue with Test 1. Test 1 is like Exercise 1 in method, and, with the practice afforded by the exercise, the pupils are able to do Test 1 without being troubled by the technic. After five minutes of working on Test 1 the pupils are directed to turn to Exercise 2 on the opposite page and are given one minute to practice that exercise. Then, by the direction of the teacher, they turn to Test 2 and continue with the same technic. In this intelligence test there are five such exercises and five such tests, each of which must be carefully timed by the examiner.

A third way of giving tests is through printed booklets containing pictures or diagrams. These are the primary or adult non-language intelligence tests. Not only must the examiner introduce the method and materials of these tests, but he must also give orally the full directions for each portion of the test.[1]

The test manual gives directions to guide the examiner in giving the test, as shown in Fig. 31. It should be noted not

[1] See Fig. 20 and Fig. 32.

only that the directions must be carefully dictated for each item but also that the time allotted for the items differs and the time for each must be carefully checked.

By a fourth method in which scales are published, the tests may be used directly by the pupils or the teacher. Such a scale as the Thorndike Handwriting Scale can be posted on the wall by the teacher and consulted by either the pupils or the teacher at will. The method is to secure a sample of handwriting. This was explained on page 104 as the selection of a sample in the middle of a paragraph which represents neither the most careful writing of the pupil nor that which is unduly hurried, and to compare that sample with the scale to determine with what scale samples it most nearly compares in quality. Pupils can be taught to do this for themselves in the judgment of handwriting. This is not the case in composition, where the judgments must be more mature. Fig. 4 (p. 61) shows how the samples of handwriting differ in quality.

A fifth way of giving tests is by personal contact of the examiner with each individual pupil. This is the method used in the individual intelligence tests. It is by far the most difficult and requires the greatest experience and skill of all tests which have yet been developed. Here practically every item of the tests must be dictated by the examiner, and every reaction must be noted by him. In a few instances the reactions of the pupil are written or drawn by him and not by the teacher. The care that the giving of such a test requires is shown in the directions for the Stanford Revision of the Binet-Simon Scale for the Measurement of Intelligence, of which a part is shown in Fig. 17 (p. 78).

Preparations for giving a test. Assuming that a test has been selected for giving to a group of pupils, as will be discussed in detail in Chapter XIII, p. 308, and that the tests are on hand, the first item in preparing to give the test

is the selection of a suitable time. It is wise to select a time of day that is likely to be free of interruptions on the part of principals or supervisors, and which is not planned so closely to a recess period or the closing of school as to result in restlessness on the part of pupils if for any reason the timing of the test is delayed. It is not wise to plan to give a test immediately after a recess period or a period of violent physical exercise, nor is it desirable to plan to give one immediately after a dinner hour, when some of the pupils may be sluggish. As a general rule a test may be most advantageously given after school has been in session for an hour or so in the morning, or at a similar time in the afternoon. After pupils have become adjusted to the schoolroom situation, and if they are not distracted by physical discomforts or excitement, they are in the best condition for the taking of a test.

Before the time that has been selected for giving the test, the teacher should make all necessary arrangements. The test papers, booklets, or sheets should be carefully counted to make sure that a sufficient number are on hand for all needs, and they should be placed where they can be easily found, but not where the pupils can come across them by chance. Children should never be given the opportunity to see, work on, or discuss the questions on test papers except during the testing periods. Therefore standard tests should never be returned to pupils. Another preparation that the teacher should make is to have ready for use a supply of sharpened pencils in sufficient quantity to provide each pupil with at least two. It is wise to have even a few more in reserve. It is not difficult for a child intent upon answering questions on a test to neglect the fact that a pencil point is breakable, and if fresh pencils are not available, time which is valuable to the pupil will be lost.

At the time of giving the test the teacher should see that desks are cleared of all books, papers, and distracting ele-

ments. Children should be taught not to waste time in erasing mistakes, but rather to cross them out and write in their corrections. For this reason it is a good plan for the teacher to prohibit the use of erasers. Ink should never be used except for purposes of getting handwriting samples.

Work that is on the blackboard should be erased, particularly if it is in any way like the materials of the test, such as would be true if there were spelling lists or other writing on the board during the spelling test, or problems in arithmetic during an arithmetic test.

Conduct of the testing period. As has been intimated, the testing period should be as free from interruption as possible. The room should be quiet, and the teacher calm and unhurried. It should always be remembered that achievement and ability are measured by these tests and that nothing which happens during the test can add to either of them, but that many events can detract from the ability of the pupil to express them. Visitors should be excluded, and the school principal should be notified of the time and duration of the test so that interruptions by supervisors will not occur.

Before the test the teacher should practice every item concerned with giving it as outlined in the manual of directions. Each statement later to be dictated to the pupils should be carefully read and its meaning and emphasis weighed. It should be remembered that every element included in the directions is placed there for some specific purpose and therefore that no word in them should be slighted or eliminated. There should be no variation from the printed directions. At times a teacher may feel it necessary to expand, simplify, explain, or restate certain directions as printed for the test, but he should guard most zealously against doing so. One of the elements which make a test standard is the standardization of the means, method, and content of the taking of the test itself. It should not be the purpose of the teacher to have

the children do the work of the test well at all costs, but rather to have them do it as well as they can under the limitations as imposed by the directions.

One of the most important elements in many tests is the timing of the parts of the test. The preferred way of timing a test is by means of a stop watch, which should be a part of the equipment of each school. By the use of a stop watch the examiner is freed from the necessity of watching or calculating time. If a test is limited to twenty minutes he can start the watch with one hand while holding the direction manual in the other when he utters the word "Go," and can say "Stop" when the hands of the watch reach exactly twenty minutes. When a stop watch is not available, timing can be done accurately with any watch which has a second hand. A woman's wrist watch is usually not equipped with a second hand, so that it is frequently necessary for a teacher to borrow a watch which does contain one. Even when using a watch with a second hand constant calculation and checking are necessary. The only safe means is to make a written notation of the beginning point and ending point of each item of the test. This will be much facilitated if a blank form is made up before beginning the test and is held ready for use during the test. This blank form could be arranged as in Fig. 42.

The appearance of such a time sheet at the conclusion of a test might be as illustrated in Fig. 43.

Every means by which a teacher can encourage pupils to attack a test with zest should be utilized, so long as it does not violate the directions given in the test manual and does not aid the pupils in their performance. The pupils should be encouraged to do their best, they should be kept at work if possible, and they should be given every opportunity to achieve. Direct help during the progress of a test should not be given, and children who seem to be hopelessly lost should

not be given help beyond that prescribed in the test manual. Questions as to the meaning of statements, the intent of a question, and the like should be answered by telling the pupil to do his best, and not by giving him a lead which would explain the matter to him. The teacher should *not* say "Read your directions more carefully" or "Did the directions tell you to do this?" He should not make comments,

PART 1	Time begun . . .	hr. . . .	min. . . .	sec.
	Time allowed		min. . . .	sec.
	Time to stop . .	hr. . . .	min. . . .	sec.
PART 2	Time begun . . .	hr. . . .	min. . . .	sec.
	Time allowed		min. . . .	sec.
	Time to stop . .	hr. . . .	min. . . .	sec.
PART 3	Time begun . . .	hr. . . .	min. . . .	sec.
	Time allowed		min. . . .	sec.
	Time to stop . . .	hr. . . .	min. . . .	sec.

FIG. 42. Time sheet for test

or show by facial expression either approval or disapproval, since all these elements supplement the test directions or pupils' achievement by undue emphasis. Directions should not be repeated unless the manual states that they may be repeated, or unless some unusual situation during the testing makes repetition fair. In a certain classroom near a street-car line an important direction was given at a time when the voice of the examiner was overborne by the noise of a passing car. Since it was obvious that only those pupils a few seats away from the examiner had heard the direction, repetition was legitimate. Such cases are rare, however; and where

there is doubt, it is usually wiser to do nothing than to violate the letter of the test directions.

When pupils are absent during the period of the test, or even if certain pupils who are taking the test are below their normal physical well-being, the test should be given to them at some later time and in a different form from that given at the first testing. In cases where pupils seem to do far more

PART 1	Time begun	9 hr.	42 min.	17 sec.
	Time allowed		10 min.	0 sec.
	Time to stop	9 hr.	52 min.	17 sec.
PART 2	Time begun	9 hr.	53 min.	27 sec.
	Time allowed		12 min.	0 sec.
	Time to stop	10 hr.	5 min.	27 sec.
PART 3	Time begun	10 hr.	7 min.	42 sec.
	Time allowed		7 min.	30 sec.
	Time to stop	10 hr.	15 min.	12 sec.

FIG. 43. Time sheet for test

poorly than the examiner feels that they might have done, it is always wise to consider the possibility of a re-test. When such re-tests are made, the score given to the pupil should not be the average of the scores which he made on the two tests, but rather the higher score achieved on the two.

When all the papers of the pupils have been completed they should be collected. Extra copies of the test should also be collected, and no papers should be left where any child can see them at a later time. The value of standard tests lies in their novelty, and if pupils can see and practice test elements the value of the test for them is lost.

After the test papers have been collected the teacher or examiner should take immediate opportunity to check on the necessary data for the identification of the pupil, the school, the classroom, and the like. Wherever data is incomplete or unreliable it should be filled in or corrected by the teacher. The names of the pupils should be checked; and if the test papers are to be used by anyone other than the teacher who is familiar with the names, it is a good plan to underline the surname of each child on each paper, thus making sure that later listing by persons unfamiliar with the class will be accurate. Other data in the possession of the teacher should be checked on each answer paper if it is called for. The age or date of birth, sex, class, section, the teacher's name, and the like should be checked, corrected, or rewritten to insure correct reports later.

A most important item, from the point of view of its later usage, is the date of the test itself and (more frequently than one might think) the year the test was given. This dating is of special importance in intelligence testing.

Scoring of test papers. The first step in scoring test papers is the assembling of materials. Test keys, the booklet containing the scoring directions, and the like should be collected and placed where they can be consulted. The papers should usually be placed in alphabetical order, although this is not necessary. To place papers in order before beginning to score them makes it possible to place them in the same order at a later time if by any chance they should be disarranged, and it also makes it easy to assemble the data at the conclusion of the scoring and to transfer the scores to permanent records.

Scoring should be done with colored pencils, the first scoring in blue, and the check scoring in red to distinguish it. Therefore it is a good plan to have a supply of sharpened blue and red pencils ready so that the scoring need not be delayed when it is once started.

It is also a good plan to have available some paper or perhaps a pad of paper on which notes can be made during the progress of the scoring, particularly notes as to the interpretation of doubtful elements or answers, so that at the conclusion of the scoring or in the check scoring these can have special attention.

When the papers have been assembled the actual scoring can begin. It is usually more efficient for two people to work together in scoring papers. In this case one person can dictate the answers for the test from the directions or from the test key, and the other can make the necessary checks on several papers at a time.

Only one portion of a test should be scored at a time, inasmuch as the scoring can be done more quickly and accurately in this way. When a test consists of only a single sheet of paper, as in the Woody-McCall Arithmetic Test, the entire test can be scored at one time; but when it is in the form of a booklet, only one page should be scored at a time. The procedure in this case is to score the one page, marking the errors or successes as the form of scoring may indicate, and then, when the page is completed, to add up the score for that page. When one page has been completed, it should be turned over and the paper set aside. The scorer may then pay attention to the same element scored in the first paper on the next test in the pile; or, when there is a reader and several tests are being scored at once, the next group of papers may be arranged for correction.

When all the papers have been completed, a re-check of the scoring should be made. If there are two or more persons working together, one of those who checked should replace the reader of the key, and the second check should be made with a different color of pencil. If blue was used for the initial scoring, the check should be in red. Every item scored in the first scoring should be checked against the key, and a

re-addition made of the scores. A red check should be placed beside each of the blue checks made in the first scoring; or, if an error is found, the new check should be indicated beyond the possibility of a doubt in making the new addition of the score. When the re-addition of the score for each page has been completed, the first addition, if correct, can be circled with red; if it has been found incorrect, it can be corrected by crossing out the incorrect score and placing the correct one near it.

When the re-check has been completed, the scores for each section of the test should be transferred either to the first page of the test (if there is a place there for the final addition of the scores) or else added up on the scratch pad. This final addition should itself be checked and circled with red to indicate this.

With this completed, if each step has been done carefully the scorer should have the test papers in alphabetical order, with the final score appearing on the first page of the test in some inconspicuous place and with the scores ready for interpretation.

When the scoring is interrupted or when the papers must be left for an interval, it is always well to indicate the point to which progress has been made. A marker placed in the last paper scored, a notation made on one of the pad sheets, a distinctive mark at the point where the scoring stopped, or any other device that will make it clear to the scorer when he returns to the work just where it was left off, will aid materially and will save time.

When there is a large number of scorers (as when college or normal-school students conduct the testing) these procedures become somewhat less tedious, because they need to be done only once for all the papers. Great care should be taken to secure uniform accuracy in the work of the scorers. The papers can be distributed among the scorers so that each

scorer has several and so that all the papers are being scored at once. The reader then reads, item by item, the proper answers to the questions, and the scorers work on the papers which they hold. When the first scoring has been completed, the papers can be exchanged and the reader can then go through the process a second time. When there are questions regarding the pupil's intent in making his answer or the exact score to be given to an answer which is not clear, such questions should be decided in every case by the reader so that all the papers may have exactly the same interpretation.

Assembling results. As will be shown in later chapters, the raw scores that are received by the pupils must be changed into other forms before they can be properly interpreted. Since the procedures for changing these scores are in the main mechanical, and since the more mechanical they can be made the more accurate they are likely to be, the teacher or examiner should decide, before attempting to assemble the results, just what purposes the test is to serve, and should make preparations accordingly. It is to be presumed that the teacher has made this decision before the tests were even selected. At the conclusion of the scoring, however, these purposes should be carefully planned and all preparations made for the gaining of the proper transmutations of scores, the proper assembling of data, and the complete interpretation of the scores in terms of the purposes. This preliminary planning, together with the preparations which can thereby be made, will save much time as well as energy and will prevent the necessity for doing some work more than once.

Later chapters deal with problems in connection with the transmutation of raw scores into various kinds of measures that are helpful to the teacher and with the use of these measures in diagnosing the difficulties of pupils, in teaching, supervision, classification, and the like; and each of these discussions deals fully with the special technics in the as-

sembling of data, the special types of treatment of data, the special materials necessary to find the measure to permit the special use, so that it is unnecessary to consider these matters here. It is necessary, however, to emphasize the fact that a preliminary survey of the work to be accomplished will make it possible to plan the sequence of procedures, to assemble the data in the quickest and easiest way, and to obtain the desired results with the greatest accuracy for later use.

1. Object: To find the grade level of a class in reading:
   a. Give a reading test to the class.
   b. Find the grade scores of the pupils.
   c. Find the mean of the grade scores.

2. Object: To find the rating in effort of a class in arithmetic:
   a. Give an arithmetic test.
   b. Find the T_a scores of the pupils.
   c. Give an intelligence test to the same class at nearly the same time.
   d. Find the T_i scores of the pupils.
   e. Calculate the F scores of the pupils.
   f. Find the mean of the F scores.

FIG. 44. Two lists of testing plans

Each procedure which is contemplated can be prepared for beforehand. A list should be made of the various operations to be performed, and the sequence in which they most logically follow. When this listing has been made, the steps for each computation should be listed and a sheet of paper prepared for the computing or transmutation of each item. When this has been done, the papers can be handled with greater ease and the examiner can work swiftly and surely without doing unnecessary work or leaving undone that which is needed. Fig. 44 shows two listings of operations.

Fig. 45 shows a list of steps for each of the operations in the two lists and gives the references to the necessary data lists to be found in later chapters.

1. Object: To find the grade level of a class in reading.
   Step 1. Give a reading test (see page 119).
   Step 2. Score the papers (see page 128).
   Step 3. Construct a raw-score-grade table for the test.
      (Step 1) Determine the grade norms for the test (see page 203).
      (Step 2) Prepare a transmutation sheet (see page 204).
      (Step 3) Place the norm positions on the transmutation sheet (see page 205).
      (Step 4) Draw the norm line through the norm positions (see page 208).
      (Step 5) Make up a raw-score-grade table (see page 209).
   Step 4. Transmute the raw scores of pupils' papers to the grade scores (see page 203).
   Step 5. Find the sum of the grade scores and divide the result by the number of pupils (see page 181).
   The result of this series of steps is a mean G score which indicates the grade level of the class in reading.
2. Object: To find the rating in effort of a class in arithmetic.
   Step 1. Give an arithmetic test (see page 119).
   Step 2. Score the arithmetic papers of the pupils (see page 128).
   Step 3. Make out an arithmetic T table for the test (see page 240).
      (Step 1) Construct a raw-score-grade table for the test (see page 203).
         [Step 1] Determine the grade norms for the test (see page 203).
         [Step 2] Prepare a transmutation sheet (see page 204).
         [Step 3] Place the norm position on the transmutation sheet (see page 205).
         [Step 4] Draw the norm line through the norm positions (see page 208).
         [Step 5] Make up a raw-score-grade table for the test (see page 209).
      (Step 2) Transmute the G scores into T scores by means of Table XXXI (see page 243).
   Step 4. Transmute the arithmetic raw scores into T_a scores by means of the arithmetic T table made in the previous step.
   Step 5. Give an intelligence test to the pupils (see page 252).
   Step 6. Score the intelligence test (see page 254).

FIG. 45. List of steps for the testing plans of Fig. 44

Step 7. Make out a T table for the intelligence test (see page 261).
(Step 1) Construct a raw-score-age table (see page 212).
[Step 1] Determine the age norms for the test (see page 212).
[Step 2] Prepare a transmutation sheet (see page 213).
[Step 3] Place the norm positions on the transmutation sheet (see page 214).
[Step 4] Draw the norm line through the norm positions (see page 215).
[Step 5] Make up a raw-score-age table for the test (see page 216).

or

[Step 1] Find a complete raw-score-age table in the intelligence-test direction manual (see page 258).
(Step 2) Transmute raw scores to T scores by means of Table XXXI (see page 245).
Step 8. Transmute the intelligence-test raw scores into T_i scores by means of the intelligence-test T table made above.
Step 9. Calculate the F_a scores, using the formula

$$[T_a + 50] - T_i = F_a \text{ (see page 298).}$$

Step 10. Find the sum of the F scores and divide by the number of pupils (see page 426).

or

Step 9. Find the sum of the T_a scores and divide by the number of pupils (see page 427).
Step 10. Find the sum of the T_i scores and divide by the number of pupils (see page 427).
Step 11. Calculate the mean F_a score according to the formula

$$[MnT_a + 50] - MnT_i = MnF_a.$$

The result of this series of steps is a mean F score in arithmetic which indicates the effort rating of the class as a whole.

FIG. 45 (Continued). List of steps for the testing plans of Fig. 44

A prime essential is to keep all records as neatly as possible. Probably no element in statistical procedure is so helpful as neatness in the accumulation and listing of results. Another helpful element is to keep all temporary data and calculations on sheets which may be discarded as soon as they have served their purpose. Data of permanent value should be

kept, of course, but calculations and other temporary data are useless and need not be kept after their results have been tabulated or used.

Testing a school. Giving tests to all the grades in a school involves only a few more procedures and preparations than does the giving of tests to a single classroom. If it is assumed that a group of tests has been selected to be given to a school, and that the tests have been procured, the procedure might be somewhat as follows:

Some person should be selected to be the director of the testing. He should be one who is familiar with the giving of standard tests and the treatment of results and who is competent to advise with the persons who are to do the actual examining. His work is to assemble the tests to be given, to appoint the hours and days during which they are to be given, to assemble a sufficient quantity of test directions and keys, to divide the tests into the proper quantities for the several rooms, to hold meetings with the teachers or examiners for the purpose of discussing the problems involved, to distribute the test material and to check its return, and to supervise the scoring. It is important that he should be familiar with testing and that he should hold the confidence of the examiners.

If the testing is to be done under the usual conditions of school practice and not by some outside agency, it is probable that the examiners must be the teachers of the schoolrooms and that the director of the testing must be the principal of the school or some teacher within the school who has made a study of testing.

The form of approach to the work will vary according to the circumstances in each school. To be valuable, every effort should be made to enlist the active effort and coöperation of every teacher. This is not always easy; but teachers are, as a rule, anxious to be helped in their teaching and are

eager to do anything that will help them to do their work better. S. S. Brooks has shown in his book *Improving Schools by Standardized Tests*[1] how he accomplished this end with a group of teachers with whom he worked. His descriptions should be an inspiration to any leader who desires to concentrate the work of his teachers on the giving and using of standard tests.

The next step is to call the teachers together for the purpose of discussing the problems involved in the testing program. The types of tests should be placed in the hands of the teachers, and a tentative time schedule should be planned so that they can make arrangements to be free at the time decided on. When a single test is to be given in all grades, it should be planned to give the test to all the pupils at the same time, and the teachers should be given ample time to make ready for it. It is not necessary, and it is rarely desirable, to have the pupils know that tests are in prospect.

After the determination of the testing periods the teachers together should carefully go over the directions given in the testing manual, and the director should at the same time emphasize any important points with reference to giving the tests. Such items as have been given in this chapter are appropriate for a discussion of this kind, particularly if the teachers are relatively unfamiliar with testing and are not aware of the great necessity for following the directions minutely and the time allotments exactly. At the conclusion of this conference each teacher should know exactly what he is expected to do and also what elements need further study on his part. The teachers should take samples of the tests to be given and also copies of the direction manuals. They should study these carefully and practice on them before attempting to give the tests to the pupils. It is always well to have another meeting of the examining group at a later time,

[1] Houghton Mifflin Company, Boston, 1922.

so that any questions relative to procedures and practices can be solved before the testing period begins.

The director should make out a chart of the testing, to be distributed among the teachers. This chart should list the procedures which the teachers are expected to follow, so that

1. Count the booklets or papers in your package.
2. Distribute one to each child.
3. Collect booklets or papers.
4. Check complete information as to name, sex, age, etc. from teacher's class book.
5. Get teacher's promotion rankings (at the time of intelligence test only).
6. Insert completed papers in envelope and tie or seal.
7. Mark plainly on the outside:
   (1) Number of papers
   (2) Grade
   (3) School
   (4) Teacher
   (5) Names of absent pupils
8. Keep unused papers separate from others and deliver all materials to office promptly.

FIG. 46. Step list

as each item is completed it can be checked off. Such a chart for the use of the individual examiner (particularly for teachers who are not experienced in giving tests) would help to achieve uniformity in the results of the testing (see Fig. 46).

The director should also make out a chart of the general program. This chart should show the number of rooms to be tested, the number of children in each room, the amount of

supplies plus extras needed for each room, the name of the teacher or examiner in each room, and like data. This sheet should be used as a progress sheet during the testing. The tests should be placed in groups sufficient for each of the rooms listed on the progress sheet. When a sufficient supply of tests has been counted out and bundled for the room, a check mark should be placed on the progress sheet for that room. When the bundle of tests is delivered to the examiner before the test period, the check can be crossed, indicating that a counted and checked number of tests has been delivered. On the return of the bundles of completed tests to the director at the conclusion of the testing, these crosses should be circled, indicating the return and check of the completed papers. In this way a director is enabled to see any variation in the proceedings and can rectify a mistake at once. A sample of such a progress sheet is given in Fig. 47.

When any particular test has been given, the director should be empowered to call the teachers together for the purpose of scoring the tests. It is not necessary for a teacher to score the papers of the pupils of his own room; so the papers may be distributed in chance order. Each teacher should work on not more than five papers at one time, with the director leading in the scoring. With five papers spread on a desk, and with the leader slowly giving the answers to each item, the answers on the five papers can be checked in order. When the first five papers are completed on one page the next five can be done in order, and the process repeated until all the papers have been scored. It is a tedious process to check the scoring, but it is always wise to do so to make sure that the original scores which are used in the judgment of any pupil are as accurate as possible.

When more than one test is given, it is possible to give the tests in any way that is convenient, provided only that the pupils are not fatigued to the extent of boredom and careless-

School	Grade I	Grade II
.....	Teacher.....	Teacher .....
	No. Pupils.....	No. Pupils .....
Date.....	Papers out ...	Papers out ...
Hour.....	Papers in	Papers in
	Completed...	Completed...
Subject.....	Unused ...	Unused ...
Test.....	Spoiled ...	Spoiled ...
Form.....	Returned ...	Returned ...

Grade III	Grade IV	Grade V
Teacher.....	Teacher.....	Teacher .....
No. Pupils.....	No. Pupils.....	No. Pupils .....
Papers out ...	Papers out ...	Papers out ...
Papers in	Papers in	Papers in
Completed...	Completed...	Completed...
Unused ...	Unused ...	Unused ...
Spoiled ...	Spoiled ...	Spoiled ...
Returned ...	Returned ...	Returned ...

Grade VI	Grade VII	Grade VIII
Teacher.....	Teacher.....	Teacher .....
No. Pupils.....	No. Pupils.....	No. Pupils .....
Papers out ...	Papers out ...	Papers out ...
Papers in	Papers in	Papers in
Completed...	Completed...	Completed...
Unused ...	Unused ...	Unused ...
Spoiled ...	Spoiled ...	Spoiled ...
Returned ...	Returned ...	Returned ...

FIG. 47. Progress sheet

ness. When tests are given in a general school survey by college students, it is frequently necessary that they be crowded into a short time; when they are given by the teachers in a school, it is possible to distribute the testing over a longer period.

The question is frequently raised as to whether a teacher should test his own class. It is true that a teacher is interested in the work of his own pupils and is desirous of having them do as well as possible. There is no handicap unless the teacher is led to violate some of the essential rulings in giving the test. It is usually an asset to have the teacher of the group of pupils conduct the testing with them. It gives the pupils a confidence they might not otherwise have, and it frequently provides an incentive for the work. However, if the teacher's zeal should lead him to give such help to his pupils as is unauthorized, it would be far better to have a stranger conduct the testing.

Teachers sometimes feel that a school-wide test is likely to result in a criticism of themselves rather than of their pupils. Such a feeling is unreasonable and wrong. If the pupils are not being taught what they should know, if they are not progressing as they should, and if there is any way in which their progress or development can be accelerated or improved, it is the duty of the teacher to use such means as are at hand to know the situation. No good can result from not knowing it, and no just criticism can be directed at the teacher who is brave enough to seek the truth.

Chapter summary. Tests may be given in a variety of ways. Some are so arranged as to be given by dictation by the examiner; others are in the form of printed booklets to be filled out by the pupils and passed in to the examiner; others, usually non-language ability or intelligence tests for illiterate adults or children in the primary grades, consist of printed booklets of pictures to be filled in or marked according to the examiner's directions. There are also the printed scales,

for use either by the teacher or by the pupils; and lastly such tests as require the contact of the examiner with the individual pupil, as in the case of the individual intelligence tests.

In giving a test the first item to decide is the day and time of day when the test is to be given. When this has been selected, provision should be made for excluding visitors, supervisors, and others who might interrupt and thus lower the level of the pupils' achievement. The teacher should prepare for the test by making sure that there are sufficient test papers on hand, by having all distracting materials removed, by preparing a sufficient quantity of sharpened pencils, and by removing all related types of work from the blackboards and bulletin boards.

During the testing period there should be as little distraction as possible. The teacher should have practiced the giving of directions so as to be able to give them fluently and clearly, and should follow the directions of the test manual implicitly. Time should be carefully checked, and a time sheet kept during the progress of the test will be a material help. A record should be kept of the pupils who are absent during the testing period, so that another form of the test can be given to them at a later date.

When the testing has been completed, all the papers should be collected immediately and all unused papers checked against the original supply. No pupil should be allowed to have a paper after the conclusion of the test. When the papers have been collected the teacher or examiner should check the important information concerning the identification of the pupils and should see to it that this information is complete for every paper.

The first step in scoring the papers is to assemble and arrange the papers so that the scoring can take place with the least amount of effort and the least possibility of error. The scoring should proceed item by item and page by page

until completed. At the conclusion of the first scoring it is good practice to re-score in order to be absolutely fair to each pupil.

When the scoring is completed, the results should be assembled in terms of the later use to which the scores are to be put. Each contemplated procedure can be planned for and much useless labor avoided by making a careful preliminary plan of procedures. It is trite but necessary to emphasize the importance of neatness and order in the accumulation and transfer of data. Care in this respect decreases the possibility of error and increases usefulness of the results.

When testing a school some one person should be selected to direct the work. This person is responsible for all the many details of organization, checking, and assembling the final results. Each teacher should have a check sheet to use in the direction of processes, and the director should keep a carefully worked-out progress chart so as to keep in touch with the stage of the testing process at all times.

When a test or a group of tests has been given, it is the director's duty to supervise the scoring and to complete all the details in connection with the after-scoring interpretations.

There is no reason why earnest and coöperative teachers cannot combine to give tests to their pupils and to receive from those tests great help in the conduct of their work and in their understanding of their pupils.

SELECTED BIBLIOGRAPHY

BROOKS, S. S. Improving Schools by Standardized Tests, chaps. ii–iv. Houghton Mifflin Company, 1922.

GREENE, H. A., and JORGENSEN, A. N. The Use and Interpretation of Educational Tests, chap. iv. Longmans, Green & Co., New York, 1929.

MCCALL, W. A. How to Measure in Education, pp. 227–233. The Macmillan Company, 1922.

MONROE, W. S., DEVOSS, J. C., and KELLY, F. J. Educational Tests and Measurements (Revised and Enlarged Edition), chaps. x, xii. Houghton Mifflin Company, Boston, 1924.

PART III. MEASURES

CHAPTER VII

MEASURES OF MASS

Meaninglessness of raw scores. When a test has been given and corrected, each paper in the group is given what is termed a "raw score." This is the total number of points that each pupil has achieved in the test. No one of these raw scores, however, is meaningful by itself, even when taken in conjunction with the total number of points which any pupil might possibly have been able to get. Thus, if the test were one for which the raw score, if every part had been done correctly, would be 120, and if some pupil received a raw score of 60, it would not mean that he had a score of 50 per cent, or necessarily anything like that: it would simply mean that out of a possible 120 points he had achieved 60. The surface indications of this score do not reveal whether he has done well or poorly, whether he has done more or less than one could expect of him. There is no surface indication of anything, save that he made a raw score of 60. Until something more is done, until some interpretation is made of the score of 60, the raw score will continue to be meaningless.

To make a raw score meaningful, it must be changed by some orderly process into some different form that does have meaning. The whole process of standard testing and the special technics that have been developed in this field are those which enable one to translate the raw scores into other forms of scores that are meaningful. No one score into which

a raw score can be changed can include within it all the useful meanings which a teacher may need for different purposes, and many different scores have been developed for various purposes. This section is devoted to a description of the more important measures into which raw scores can be turned or which can be obtained from raw scores, with directions to enable the teacher to find them and use them. Later sections are devoted to the uses that teachers can make of combinations of these measures for particular purposes.

Types of measures most in use. Although there are many statistical measures that might be used by teachers for the interpretation of raw scores, there are a few of much significance with which all teachers who handle standard tests should be familiar.

1. *Measures of mass.* For any group of raw scores there are in general use measures which are called measures of mass. These are measures that show the relationship of the entire mass of scores under consideration. For a measure that will show whether a group of scores is widely distributed or narrowly distributed, and how widely or how narrowly, and that will show whether there are peculiarities in the distribution which would have an effect on the interpretation of any of the scores (of the test or of the pupils as a whole), a measure of mass is used. Those in most frequent use are the order distribution, the rank distribution, the frequency distribution, and the frequency surface.

2. *Measures of points.* Measures of points describe the character of a group of scores by indicating some point within them as a criterion. Such measures are the mode, the median, the first quartile point, the third quartile point, and the arithmetic mean. There are many other point measures, as well as measures of relationship, that are in statistical use, but for most of the teacher's purposes those that have been mentioned are sufficient.

3. *Measures of achievement.* From the point of view of the teacher measures of achievement assume a great importance, since it is from these measures that he can get some judgment as to the actual achievement of his pupils and can compare in achievement his pupils with those of other teachers. The measures of achievement in general use are the grade scores, the age scores, and the T scores. These scores taken individually for any class of pupils reveal for each pupil the actual achievement which has been made; when combined as measures of mass they give an interpretation of the work of an entire group.

4. *Measures of ability.* Inasmuch as pupils differ greatly in their ability to learn and in their reaction to learning, measures of ability form an important group. These measures, the T score in ability and the mental age, are becoming increasingly more useful to teachers.

5. *Measure of rate.* Measures of achievement and of ability show the status of a pupil at any given time; but they do not indicate the rate at which he can learn if he uses his full ability, the rate at which he has learned, or the rate at which he might learn. These measures are known as the intelligence quotient, the achievement quotient, the promotion quotient, the B score in intelligence, the B score in education, and the B score in promotion.

6. *Measures of effort.* These are in reality combinations of measures of achievement and measures of ability and indicate how nearly pupils have done as much as they might have done, in consideration of the ability with which they are endowed. Measures of effort, the achievement ratio and the F score, are relatively new in education, but they give for the first time a means whereby a teacher can judge the achievement of a pupil without the necessity of comparing him with other pupils and without necessarily taking into consideration his standing in the class.

MEASURES OF MASS

Measures of mass give a variety of pictures of the status of a group, the character of the picture being dependent on the type of measure in use. In the order distribution the picture is merely a sketch, in the rank distribution it is more similar to a line diagram, in the frequency distribution it is something like a drawing, and in the frequency surface it has the characteristics of a photograph. According to the type of description of the group which is desired, the teacher should choose the form of the mass measure that will be of greatest use.

The order distribution. The order distribution is one in which all the scores are placed in the order of their size. This is the simplest of all the measures of mass, and is frequently the one used because it serves as a preliminary to many other procedures.

Beginning with the best score on the test in question for any group, list the scores in the order of their size, including the scores of every pupil even though there is more than one score of the same size. In the following set of raw scores, which will be used in all the operations to follow, the order distribution is made by first picking out and setting down the best score, 112. The next best score among the raw scores, 109, is then set down, and this is followed by 107, 106, and 104. It will be noted that there are two scores of 104. They should both be set down. The same procedure is followed for the remaining scores. Table I shows the raw scores as they were taken directly from the pupils' papers; Table II shows the same set of scores placed in order distribution.

The rank distribution. The rank distribution shows something similar to the order distribution, but carries the meaning further and shows the character of the distribution more

TABLE I. SET OF RAW SCORES

Taken from the Test Papers of a Group of Pupils

55	79	35	82	76
94	64	82	92	71
57	100	78	68	73
41	66	92	84	61
92	43	104	76	75
73	56	50	104	72
56	93	64	63	70
67	75	97	95	112
103	54	68	71	75
56	45	48	80	87
52	93	91	74	109
81	107	100	96	86
98	51	77	62	59
51	83	73	71	74
79	87	65	77	59
88	106	66		

TABLE II. ORDER DISTRIBUTION OF RAW SCORES AS IN TABLE I

112	97	88	79	74	68	61	51
109	96	87	79	74	68	59	51
107	95	87	78	73	67	59	50
106	94	86	77	73	66	57	48
104	93	84	77	73	66	56	45
104	93	83	76	72	65	56	43
103	92	82	76	71	64	56	41
100	92	82	75	71	64	55	35
100	92	81	75	71	63	54	
98	91	80	75	70	62	52	

clearly. It gives the ranking, as among all the scores, of any one of them in terms of all the rest. It is found as follows:

STEP 1. *Making an order distribution*

The first step in finding a rank distribution is to set down all the raw scores as they are taken from the papers in an

order distribution, as has been described above and illustrated in Table II.

STEP 2. *Making a numerical order*

The second step is to make, in a column of scores in the order distribution, a numerical rating for each score, giving each a number, the best score having the number 1. Thus the numerical rank score 112 would be number 1, the score 109 number 2, the score 107 number 3, the score 106 number 4, the two scores 104 numbers 5 and 6, and so on, until all the scores have been numbered.

STEP 3. *Finding the rank of each score*

In a third column can now be filled in the ranking of each size of score. This is done by giving each score a rank equivalent to its numerical rank. When two or more scores are the same size the ranking is the average of the numerical ranks of them all. Thus, the score 112 has a rank of 1, 109 a rank of 2, 107 a rank of 3, and 106 a rank of 4; the scores of 104 have numerical ranks of 5 and 6, the average of which is their ranking, 5.5; the score of 103 has a rank of 7, and for the score of 100 the average of the numerical ranks, 8 and 9, is 8.5. The same procedure is carried out for all the scores, and the completed table is a rank distribution, as is shown in the third column of Table III.

The rank distribution is used when it is wished to find the relative placement of any score within a group of scores and to know its relationship to all of them. Thus, the pupil with a score of 78 has a ranking of 33 out of a class of 78 pupils, although there are actually only twenty-five sizes of score that exceed his.

The frequency distribution. In the frequency distribution the scores are grouped in a table where the size of the groups is in relation to the number of raw-score units of the test and

TABLE III. RANK DISTRIBUTION OF RAW SCORES AS GIVEN IN TABLES I–II

SCORE	ORDER	RANK	SCORE	ORDER	RANK
112	1	1	74	41	
109	2	2	74	42	41.5
107	3	3	73	43	
106	4	4	73	44	44
104	5	5.5	73	45	
104	6		72	46	46
103	7	7	71	47	
100	8	8.5	71	48	48
100	9		71	49	
98	10	10	70	50	50
97	11	11	68	51	51.5
96	12	12	68	52	
95	13	13	67	53	53
94	14	14	66	54	54.5
93	15	15.5	66	55	
93	16		65	56	56
92	17	18	64	57	57.5
92	18		64	58	
92	19		63	59	59
91	20	20	62	60	60
88	21	21	61	61	61
87	22	22.5	59	62	62.5
87	23		59	63	
86	24	24	57	64	64
84	25	25	56	65	66
83	26	26	56	66	
82	27	27.5	56	67	
82	28		55	68	68
81	29	29	54	69	69
80	30	30	52	70	70
79	31	31.5	51	71	71.5
79	32		51	72	
78	33	33	50	73	73
77	34	34.5	48	74	74
77	35		45	75	75
76	36	36.5	43	76	76
76	37		41	77	77
75	38	39	35	78	78
75	39				
75	40				

where the table shows the number of raw scores actually achieved by the class in each of the groups.

STEP 1. *Determining the number of test units in each score group*

The first step is to determine the number of test units which should be placed in each score group of the frequency-distribution table. This is commonly called the step interval of the distribution; if the distribution is to be meaningful it is important that the step interval be right. There are two ways of determining the size of the step interval: one with groups of a hundred or less, and the other with groups of more than a hundred.

When there are less than a hundred scores the size of the step interval for the frequency distribution is determined by locating the highest and lowest scores in the group, subtracting the lowest from the highest, and dividing the result by the figure found in Table IV. The result is the number of score units to be placed in each score group. The formula is

$$\frac{\text{Highest score} - \text{lowest score}}{\text{(Divisor from Table IV)}} = \text{step interval.}$$

In the set of raw scores as given in Table I the highest score is 112 and the lowest is 35. This is the test range. There are 78 scores. Table IV states that when the number of pupils is between 65 and 79, the divisor 15 should be used. Dividing the test range, 77, by 15 gives 5.1.

$$\frac{112 - 35}{15} = \frac{77}{15} = 5.1.$$

Since 5.1 scores cannot be placed in a single score group, the nearest whole number of scores must be used — in this case 5. This gives the number of test units to be placed in each score group in the distribution table. When the scores are found in tenths then each will be a discrete unit.

TABLE IV. DIVISORS TO BE USED IN FINDING NUMBER OF UNITS FOR SCORE GROUPS IN FREQUENCY DISTRIBUTION OR FREQUENCY SURFACE ACCORDING TO THE NUMBER OF PUPILS TESTED

When the number of pupils is 16 or less, divide range by 7
When the number of pupils is from 16 to 20, divide range by 8
When the number of pupils is from 21 to 25, divide range by 9
When the number of pupils is from 26 to 30, divide range by 10
When the number of pupils is from 31 to 36, divide range by 11
When the number of pupils is from 37 to 42, divide range by 12
When the number of pupils is from 43 to 50, divide range by 13
When the number of pupils is from 51 to 64, divide range by 14
When the number of pupils is from 65 to 79, divide range by 15
When the number of pupils is from 80 to 100, divide range by 16

When the number of pupils is greater than a hundred the number of units is determined in the following way: The test range should be determined. This is done as shown above, by subtracting the lowest score from the highest, giving the range of the test in score units. This should be divided first by 10 and then by 20, and the most convenient whole number between the two quotients should be selected as the number of score units to place in each score group. Thus, if there were 255 pupils, with the highest score on the test 97 and the lowest score 32, the step interval would be determined as follows:

$$\frac{\text{Highest score} - \text{lowest score}}{10} = \text{greatest number of units.}$$

$$\frac{\text{Highest score} - \text{lowest score}}{20} = \text{least number of units.}$$

$$\frac{97 - 32}{10} = 6.5.$$

$$\frac{97 - 32}{20} = 3.25.$$

In this case either 4, 5, or 6 units could be accurately placed in each score group. It is probable that the most convenient number of units in this case would be 5.

STEP 2. *Determining the beginning point of the step interval*

When the number of score units in each step interval has been determined, it then becomes necessary to find the exact beginning and ending points of the step intervals, as this is of great importance in using the frequency-distribution table. The beginning point of the step interval varies according to the character of the score on the test, and the character of the score is determined by the type of test which has been used. It is necessary, therefore, to inquire into the exact meaning of the scores that have been received by the pupils.

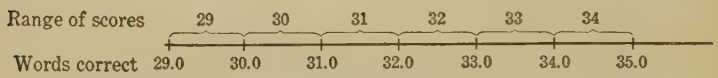


FIG. 48. The meaning of scores in a spelling test

If the test had been a spelling test, in which the number of words correct determined the score, each score would indicate the highest number of *complete* words correct, and a score of 32 would mean that the true score of the pupil was from 32 just up to but not including 33. Fig. 48 shows the meaning of scores in a spelling test.

In this case, if there were five score units in a step interval and if one score group were to contain all the scores 30, 31, 32, 33, and 34, the beginning of the step interval would be just 30, the lowest score which could be received by a pupil, and the end point would be 34.9. . . . In practice this is written 30.0 – 34.9, and would include all the discrete scores as shown above.

If, on the other hand, the test had been a handwriting test, or a composition test, or a test in which the judgment of the teacher has been called into action in determining the score, the scores would have a different meaning. In this case the teacher would have to say that the sample was better than 29 but not so good as 31, that it was nearer 30 than 29 and

nearer 30 than 31, and that therefore the score given to the sample was 30. This same reasoning would be used in determining samples valued at 31, 32, 33, 34, etc. Fig. 49 shows the meaning of scores in a composition test.

If five of these score units were to be placed in a step interval, the lowest score, 30, would really mean from halfway between 29 and 30 to halfway between 30 and 31, and so on for each of the scores to 34, which would mean from halfway between 33 and 34 to halfway between 34 and 35. The step interval would then begin with 29.50, would run to 34.49

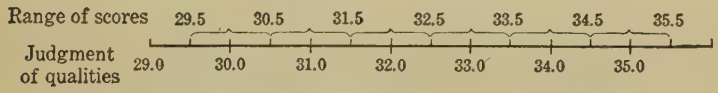


FIG. 49. The meaning of scores in a composition test

(usually written 29.50—34.49), and would include all scores between those two points. When making up a frequency-distribution table it is known that there will be no score of say 29.55 or 31.29. Nevertheless, as has been said, it is important that the theoretical beginning and ending points for the step interval should be determined, since the use of the distribution in finding other measures makes it necessary.

STEP 3. *Constructing the frequency-table tally sheet*

When the size of the step interval and the beginning points of the step intervals have been determined, it becomes possible to construct a frequency table by which to make a frequency distribution. The tally sheet for the frequency distribution should be made in neat columns so that it can be used later in finding measures of points, as described in Chapter VIII. The first column of the sheet should consist of a listing of the step intervals in terms of the beginning and ending points, as has been determined.

Beginning at the bottom of the sheet with the lowest score

and the lowest score group, the step intervals should be placed as in Table V. This table is constructed for a test in which the lowest score in each score group is the lowest score that can be obtained, as in a spelling test or an arithmetic test or an intelligence test. The lowest score is 35, so the five units in that group run from 35.0 to 39.9. The next higher group extends from 40.0 to 49.9 and so on, regardless of whether there are actually scores within the group or not. Every possible score group must be included within the table. The second column of the sheet consists of space for tallying each of the scores received by the pupils.

STEP 4. *Tallying the scores and converting the tally into a frequency distribution*

Beginning with the first score of Table I, which is 55, a mark is placed in the tally column opposite the score group which includes that score — 55.0 – 59.9. The second score, 94, is then placed in the score group 90.0 – 94.9 and is indicated by a single mark. The remaining scores are then placed in order, a single stroke indicating each score as it is tallied and the fifth stroke in each group being crossed on the previous four (𝍸) to make the final counting of the scores easier.

Table V shows the tally sheet before beginning the tally; Table VI shows the sheet at the conclusion of the first ten tallies; Table VII shows the completed tally, with the final count of tallies placed in a third column on the sheet. The final count, which merely turns the total of tally strokes in each group into a number, is the frequency distribution.

The frequency distribution is used as a preliminary operation in finding point measures. Unless a more exact picture is desired, the frequency distribution gives a good idea of the character of the work, since it shows the exact range of the test in points and the relative groupings of the pupils in the various step intervals.

TABLE V. TALLY SHEET FOR DISTRIBUTION SHOWING SCORE GROUPS AND SPACE FOR TALLY STROKES

SCORE GROUPS	TALLY SPACE
110.0–114.9	
105.0–109.9	
100.0–104.9	
95.0– 99.9	
90.0– 94.9	
85.0– 89.9	
80.0– 84.9	
75.0– 79.9	
70.0– 74.9	
65.0– 69.9	
60.0– 64.9	
55.0– 59.9	
50.0– 54.9	
45.0– 49.9	
40.0– 44.9	
35.0– 39.9	

TABLE VI. TALLY SHEET FOR FREQUENCY DISTRIBUTION AFTER TALLYING THE FIRST TEN SCORES

SCORE GROUPS	TALLY SPACE
110.0–114.9	
105.0–109.9	
100.0–104.9	/
95.0– 99.9	
90.0– 94.9	//
85.0– 89.9	
80.0– 84.9	
75.0– 79.9	
70.0– 74.9	/
65.0– 69.9	/
60.0– 64.9	
55.0– 59.9	////
50.0– 54.9	
45.0– 49.9	
40.0– 44.9	/
35.0– 39.9	

TABLE VII. TALLY SHEET OF FREQUENCY DISTRIBUTION AFTER TALLYING ALL SCORES AND CONVERTING TALLIES INTO FREQUENCY DISTRIBUTION

SCORE GROUPS	TALLIES	FREQUENCY
110.0–114.9	/	1
105.0–109.9	///	3
100.0–104.9	𝍸	5
95.0– 99.9	////	4
90.0– 94.9	𝍸 //	7
85.0– 89.9	////	4
80.0– 84.9	𝍸 /	6
75.0– 79.9	𝍸 𝍸	10
70.0– 74.9	𝍸 𝍸	10
65.0– 69.9	𝍸 /	6
60.0– 64.9	𝍸	5
55.0– 59.9	𝍸 //	7
50.0– 54.9	𝍸	5
45.0– 49.9	//	2
40.0– 44.9	//	2
35.0– 39.9	/	1

The frequency surface. The frequency surface tells at a glance the character of the work of the class. It may be constructed in two ways: from a frequency distribution or, directly, from the raw scores. The object in making the frequency surface is to show graphically (that is, by the size of space given to each score) how the various scores are distributed. A casual inspection will show the character of the surface; a careful analysis will reveal its meaning.

1. *Making the frequency surface from the frequency distribution.* When a frequency surface is to be made after a frequency distribution has been completed, the procedure is as follows:

STEP 1. *Making a frequency distribution*

The frequency distribution is made by placing all the scores on a tally sheet in score groups and finding the numerical tally of the scores as was shown on pages 150–156.

STEP 2. *Preparing a graph sheet*

The next step is to prepare a graph sheet for the frequency surface. This sheet contains squared lines and can be procured at any stationery store, or, for ordinary purposes, can be made by ruling a sheet of paper as it is needed. Quarter-inch squares to half-inch squares are suitable

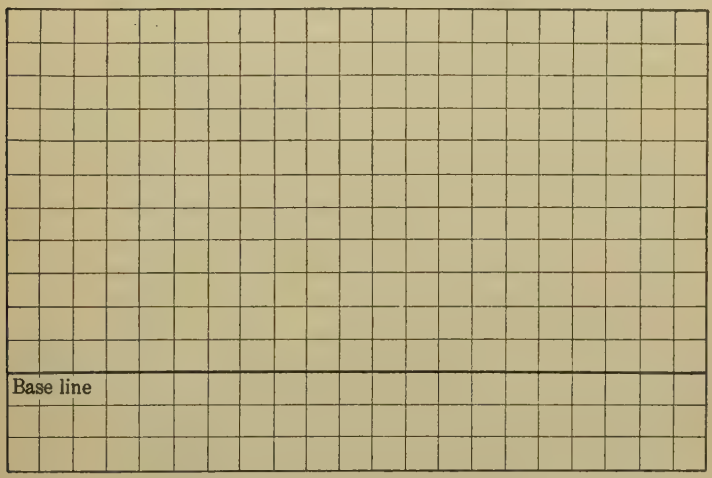


FIG. 50. Graph sheet with base line

for most purposes, although smaller squares are necessary when large numbers of scores are to be placed on ordinary letter-size paper.

STEP 3. *Drawing a base line*

Holding the long side of the graph sheet horizontally, draw somewhere on the lower part of the sheet, within four or five squares from the bottom, a line. This is called the base line, and is used as a point of reference in constructing the surface. Fig. 50 shows a graph sheet with the base line drawn as described.

STEP 4. *Labeling the columns below the base line*

The next step is to label each vertical column of squares that crosses the base line with the group labels of the various step intervals, as determined in making the frequency distribution.[1] These should start with the lowest group on the left-hand side, not too close to the edge of the sheet, but close

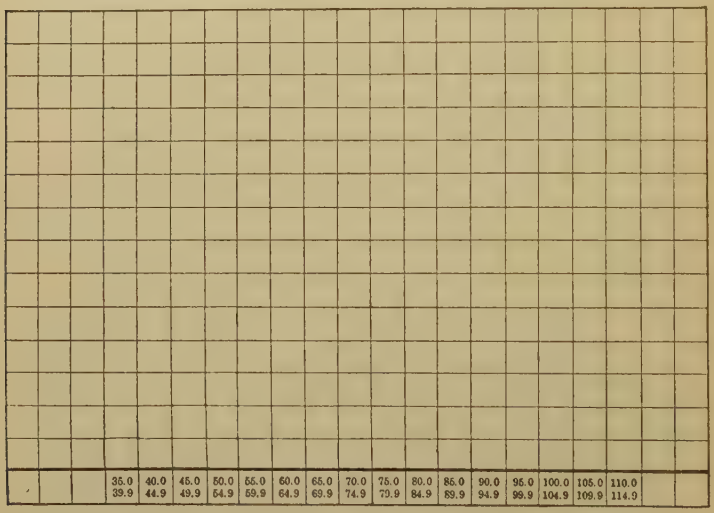


FIG. 51. Graph sheet after labeling columns below base line

enough to allow all the labels to be placed on the base line. The appearance of the graph sheet after labeling the columns below the base line is shown in Fig. 51.

STEP 5. *Placing the size of the score groups above the base line*

The final step is to place the size of the score groups above the base line in the appropriate columns, allowing one square above the base line in each column to indicate one score. Thus, in the column labeled "35.0–39.9" one square would

[1] See page 153.

be indicated. This can be done by placing a dot within the square, by outlining the square, by crosshatching it, or by indicating in some other way that it has been filled. In the next column, which is labeled "40.0–44.9," two squares above the base line would be filled in; in the third column two squares would be filled in; and in the next, which is

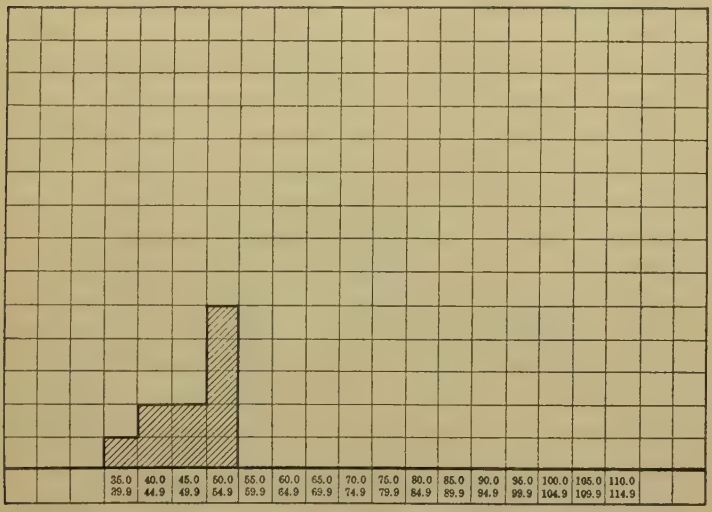


FIG. 52. Graph sheet at conclusion of filling in of four columns

labeled "50.0–54.9," five would be filled in. The appearance of the frequency surface at this point would be similar to that in Fig. 52. When all the scores have been transferred from the frequency-distribution sheet, the surface of the extreme squares should be outlined with a heavy line, and the result is the frequency surface, as shown in Fig. 53.

2. *Constructing the frequency surface directly from the raw scores.* When the frequency surface is desired and the frequency distribution has not been made, the surface can be constructed directly from the raw scores as follows:

STEP 1. *Determining the test range and computing the size of the step interval*

Just as in making a frequency distribution, in making a frequency surface from the raw score, it is necessary to determine the proper number of test units to place in each score

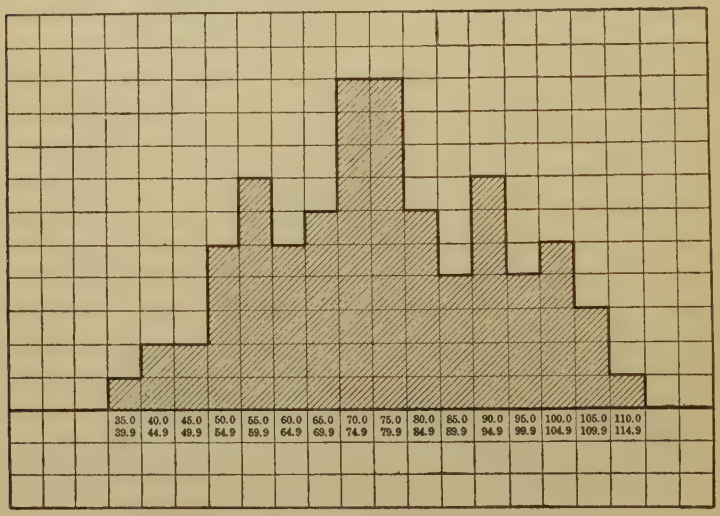


FIG. 53. Completed frequency surface

group. This is done as described on page 150. As a general rule it is undesirable to attempt to make a frequency distribution from the raw scores of a test directly when there are more than one hundred scores to deal with; when there are less than one hundred the frequency surface can be found nearly as quickly in one way as in the other, and it is simple to transmute a frequency distribution into a frequency surface or a surface into a distribution when either is needed.

STEP 2. *Preparing a graph sheet*

The second step is to prepare a graph sheet similar to that described for making a frequency surface from the distribu-

tion. This sheet may be either purchased or made, and should contain squares from one-quarter inch to one-half inch in size. Holding the sheet with its long side horizontal, draw a line near but not on the bottom of the sheet. This is called the base line, and is the point of reference for the surface or

FIG. 54. Graph sheet with base line

curve that is to be built upon it. Fig. 54 shows the appearance of the graph sheet as prepared for the labeling of the columns.

STEP 3. *Labeling the columns*

The next step is the labeling of the columns. In Step 1 the size of the step interval was determined. It now remains to place at the foot of the columns of squares which cross the base line the labels indicating the scores to be placed in each column. Starting on the left-hand side of the base line so as to leave room for the total number of columns which may be necessary to include the distribution, each column should be labeled, beginning with the column containing the lowest

score. In the example as given in Table I the lowest score is 35, and according to Step 1 the step interval has been determined as 5. This means that each of the columns is to contain all the scores actually received in five consecutive units of the test range. Thus the first column should be labeled to contain all scores of 35, 36, 37, 38, and 39, the next

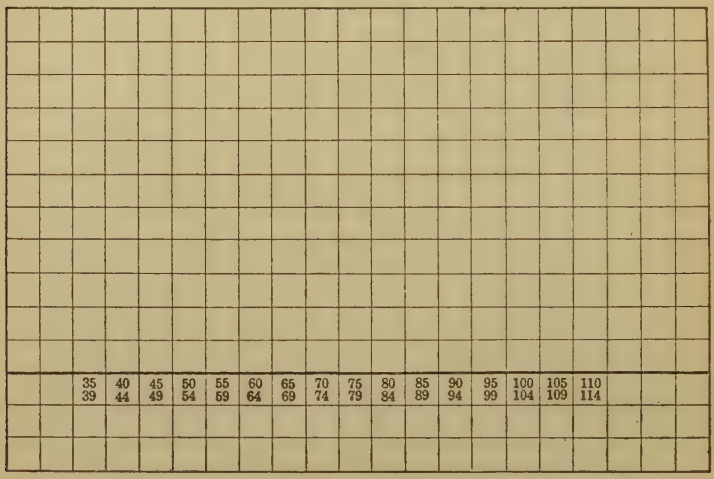


FIG. 55. Graph sheet after labeling columns below base line

column to contain all scores of 40, 41, 42, 43, and 44, and so on until the columns have been labeled to contain all the scores received by the pupils taking the test. The columns may be conveniently labeled "35–39," "40–44," and so on. Fig. 55 shows the graph sheet at the conclusion of this step.

STEP 4. *Placing the raw scores on the graph sheet*

The next step is the placing of the raw scores in their appropriate places on the graph sheet. This can be done by inserting the score itself, by crosshatching the space which indicates the score, by placing a dot in the middle of the square, or any similar convenient way. Since, when there are small

numbers of cases (as in a class of pupils), one can get a rank distribution from a frequency surface if the actual scores are given, in this case the method of indicating the scores by the actual score received will be followed. In the illustration which has been used before, the raw scores appear on the various papers of the pupils in the order given in Table I,

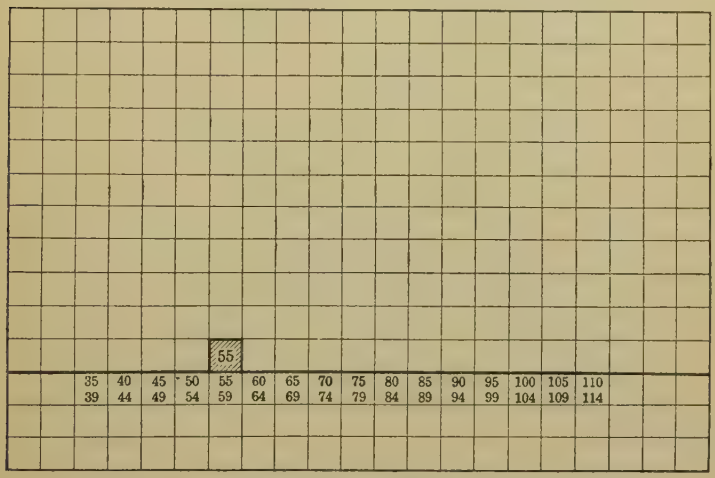


FIG. 56. Graph sheet after placement of first score

p. 147. In this group the first score is 55. This is then written in the first square above the base line in the column 55–59, which contains that score. When this has been done, the graph sheet will have the appearance of Fig. 56.

The next score found is that of 94 on the second paper. This is then written in the first space above the base line in the column labeled "90–94," which contains that score. The third score, 57, should be written in the first unoccupied square in the column labeled "55–59," which contains that score, directly above score 55. When this has been done, the graph sheet will have the appearance of Fig. 57.

All the remaining scores should then be placed in their appropriate squares and in their appropriate columns on the graph sheets until the last of the scores has been transferred from the papers to the sheet. The sheet will then have the appearance of Fig. 58. In this figure the extremities of the columns have been outlined with a heavy line. This is

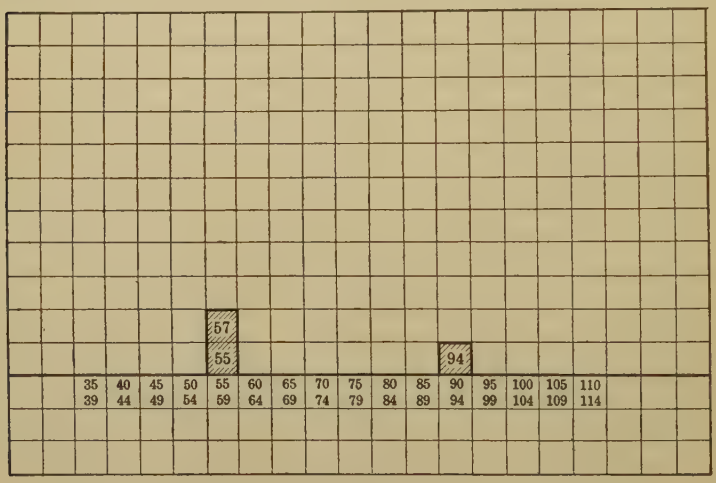


FIG. 57. Graph sheet after placement of three scores

frequently called the test curve and, as will be shown later, has great significance in the interpretation of the achievement of a class.

Use of mass measures. Mass measures in general are used mainly to interpret the character of the work which a group of pupils has done, or to give to any score a greater meaning with respect to the group of which it is a part than would be possible if the score were to be considered alone. Mass measures are necessary preliminaries to other forms of interpretation of scores and are therefore fundamental to them. Since this is the case, a discussion of the meaning and interpretation

of test curves, or frequency surfaces, will be postponed until after a discussion of the measures of points in this chapter.

Chapter summary. Raw scores are meaningless unless they are transposed into some other form. Since there are many forms of measures, each with its place in the interpretation of a score, each must be considered separately and the

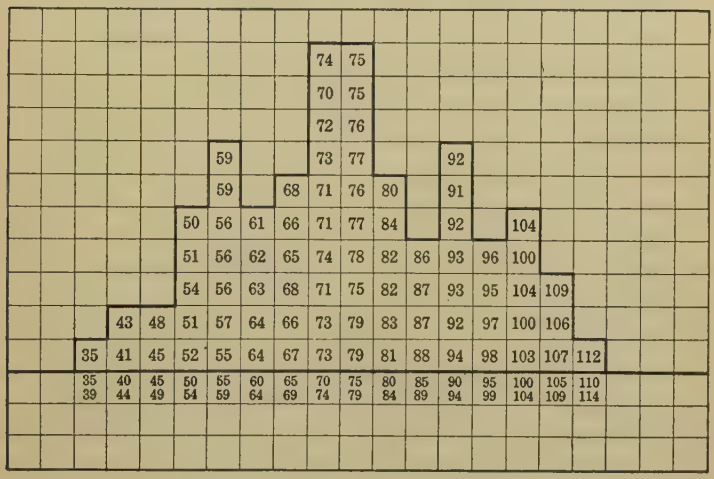


FIG. 58. Completed frequency surface

meaning and interpretation of each must be taken separately. Among the fundamental measures commonly in use into which raw scores may be changed are certain measures of mass.

The first of the measures of mass to be considered is the order distribution, which consists merely in the placement of the raw scores of a test in the order of their size.

The second measure is the rank distribution. This is found by locating the order distribution of a series of raw scores, by giving each score a numerical rating from best to poorest, and then by giving each score a ranking equivalent

to its numerical rating, which differs from the numerical rating only when there are two or more scores of the same size.

The third measure is the frequency distribution, which is more complex than the others. To find a frequency distribution it is necessary first to determine the test range by subtracting the lowest score from the highest score and then to divide the remainder by a number found in a table given, to determine the size of step intervals for the score groups. With this determination made, it is necessary to decide the meaning of the score unit and then to label a table with the various score-group labels. A tally can then be made which when completed and transposed into numerical form constitutes the frequency distribution.

The fourth measure considered is the frequency surface. This can be found either from the frequency distribution or from the raw scores. When found from the frequency distribution it is necessary to prepare a graph sheet by drawing a base line and labeling the vertical columns of cross-sectioned rulings below the base line with the appropriate group labels of the frequency distribution, beginning with the lowest group on the left-hand side of the sheet. When this is done, a square above the base line should be filled in for each score which is contained as a frequency in each score group. When the outline of the squares is emphasized with a heavy line, the result is the frequency surface.

When the frequency surface is made from the raw scores it is necessary to determine the size of the step intervals in the distribution, and from that to prepare a graph sheet with base line and column labels as previously described. After this each individual score can be taken directly from the papers of the pupils and placed in its appropriate place on the sheet. When the extremities of the columns of the sheet are outlined with a heavy line, the result is a frequency surface similar to that already described.

SELECTED BIBLIOGRAPHY

FENTON, N., and WORCESTER, D. A. An Introduction to Educational Measurements, chap. v. Ginn and Company, Boston, 1928.

GREENE, H. A., and JORGENSEN, A. N. The Use and Interpretation of Educational Tests, chap. vi. Longmans, Green & Co., New York, 1929.

GREGORY, C. A. Fundamentals of Educational Measurement, chap. ix. D. Appleton and Company, New York, 1922.

MCCALL, W. A. How to Measure in Education, chaps. xiii, xiv. The Macmillan Company, New York, 1922.

OTIS, A. S. Statistical Method in Educational Measurement, chap. iii. World Book Company, Yonkers on Hudson, 1925.

RUGG, H. O. A Primer of Graphics and Statistics. Houghton Mifflin Company, Boston 1925.

RUGG, H. O. Statistical Methods Applied to Education, chap. iv. Houghton Mifflin Company, Boston, 1917.

RUSSELL, C. Classroom Tests, chaps. x, xi. Ginn and Company, Boston, 1926.

TRABUE, M. R. Measuring Results in Education, chap. x. American Book Company, New York, 1924.

WILLIAMS, J. H. Graphic Methods in Education. Houghton Mifflin Company, Boston, 1924.

CHAPTER VIII

MEASURES OF POINTS

Characteristics of measures of points. Measures of points are in general use in the interpretation of data in many fields outside of education. They describe the character of a group of scores in a different way from measures of mass. Whereas in the case of measures of mass the entire group of scores is considered in determining the meaning of the measure, in the measures of points some one point in the group is selected as being characteristic of the entire group in some form or other. These measures may be somewhat inaccurately thought of as averages, with the different measures representing different forms or kinds of averages. The measures of points in more general use in the interpretation of educational scores are the crude mode, the median, the quartile points, and the mean. Each has its place in the interpretation of the scores, and each is of importance in terms of its peculiar values.

The crude mode. The crude mode is defined as the most frequent measure. It is calculated from the mid-point of the step which has the greatest frequency in either a frequency distribution or a frequency surface. As can be seen by reference to the frequency distribution on page 156, or to the frequency surfaces on pages 160 and 165, it is sometimes difficult to find and is unreliable. As an average score it is subject to great error. In the example given on page 156 the crude mode of the distribution of raw scores may be either of two places, since there are two score groups which have the greatest frequencies. These are the groups from 70.0 to 74.9 and from 75.0 to 79.9. Therefore the crude mode, calculated from the

mid-point of the step, may be either 72.5 or 77.5. In a rather regular distribution, when there is a most frequent case the crude mode is a useful measure. As a general rule it is better to make the effort to find a more reliable measure which will be of greater use and thus of more value.

The median. The median is a standard measure of great significance, one that has wide use and is generally understood. It may be considered as the theoretically exact middle point in a distribution of scores.

In the calculation for finding the median certain assumptions are made which may or may not be true but which must be made. It is assumed that within any score group in a distribution the scores are evenly distributed throughout the group. In the discussion concerning the exact meaning of scores[1] it was pointed out that the actual score received by a pupil in a spelling test might not be his true score, inasmuch as a score of 32 really meant that he had been able to spell thirty-two words correctly but also indicated that he could not do thirty-three words correctly, with the result that his true score might be somewhere between 32 and 33 but just not reaching 33. The calculation for the median takes this into account, and assumes that if there are two or more scores in any score group the scores do not all lie at the mid-point of the step nor at the beginning point of the step, but rather are distributed evenly throughout the step. In calculating the median from the frequency distribution shown on page 156, therefore, it would be assumed that in the step from 65.0 to 69.9 the six scores would be distributed evenly, as shown in the first half of the line in Fig. 59, and in the step next above, from 70.0 to 74.9, the ten scores there would also be evenly distributed throughout the step.

A second assumption which is made in finding the median is that because of the theoretical even distribution of scores

[1] See page 152.

throughout the score groups the calculated middle score, or median, may not have been actually received by any individual as a raw score, but is probably nearer to the true score of some pupil than the one he actually received. Placing the same idea in a different setting, it is possible to say that the median is the exact size of the theoretical middle case. This fact is shown more clearly in Fig. 60.

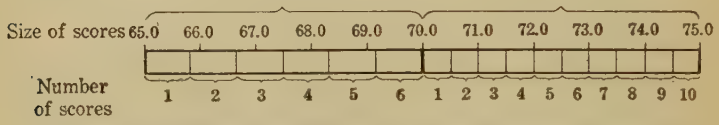


FIG. 59. Theoretical distribution of scores throughout steps in calculation of the median

Steps in finding the median. To find the median the following series of steps should be carefully followed so as to prevent confusion:

STEP 1. *Making a frequency distribution*

The frequency distribution of the scores on a test is made as described on pages 150 ff.

STEP 2. *Determining half the frequencies*

Since the median lies at the theoretical mid-point of the frequency distribution, it is necessary to determine the step and the portion of the step in which the median lies. It is necessary first to sum the frequencies and then to divide the number of the frequencies by 2. This gives the number of frequencies on each side of the mid-point. The formula is

$$\frac{N}{2} = \text{the point at which the median lies.}$$

In the illustration of frequency distribution given on page 156 the total number of frequencies is 78. Dividing this number by 2 shows that 39 frequencies lie on each side of the median.

STEP 3. *Adding the frequencies up to the step in which the median is contained*

The next step is to add the frequencies, beginning with those in the lowest step in the distribution, up to the step in which the frequency of 39, the step containing the median, is reached. In the step 35.0–39.9 the frequency is 1; in the next step it is 2, and adding this gives a total of 3; the frequency in the next step is 2 again, giving a total of 5; adding 5, the frequency in the following step, makes the total 10; the next three additions give successive totals of 17, 22, and 28; and, finally, adding the frequency in the step 70.0–74.9 gives a total of 38 $(1 + 2 + 2 + 5 + 7 + 5 + 6 + 10)$. Since the frequency in the next step is 10, it is clear that $\frac{N}{2}$ and consequently the median lie in that step.

STEP 4. *Determining the median*

The median lies at the point in the step 75.0–79.9, where the thirty-ninth score ends and the fortieth begins. This point is found by means of the formula

$$\left(\frac{\text{Remainder}}{f} \times SI\right) + \text{beginning of step} = \text{median},$$

which can be stated in words as follows: To the remainder of the frequencies necessary to reach $\frac{N}{2}$ in the step in which the median is contained, divided by the frequency of the step and multiplied by the size of the step interval, add the beginning point of that step.

The only real difficulty in the interpretation of this formula lies in the meaning of the word "remainder." Too often students believe that the word "remainder" indicates the remaining number of frequencies within the step after $\frac{N}{2}$ has been reached. *The remainder* that it is necessary to use *is the amount*, after adding up the frequencies to the step in which

the median is contained, *that still remains to reach* $\frac{N}{2}$. In the case just cited in Step 3, $\frac{N}{2}$ is 39; the addition of the frequencies up to the step in which $\frac{N}{2}$ is contained is 38; subtracting 38 from 39 leaves a remainder of 1 to reach $\frac{N}{2}$.

In the foregoing example the rest of the substitution is as follows: 1 divided by the number of frequencies in the step in which the median is contained is $\frac{1}{10}$, or 0.1; 0.1 multiplied

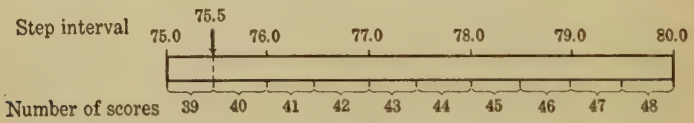


FIG. 60. Graphic representation of placement of median

by the size of the step interval, 5, is 0.5; adding 0.5 to the beginning point of the step, 75.0, gives a median of 75.5. The substitution in the formula is

$$(\tfrac{1}{10} \times 5) + 75.0 = 75.5.$$

The remainder to reach $\frac{N}{2}$ is 1, the frequency is 10, the size of the step interval is 5, and the beginning point of the step in which $\frac{N}{2}$ is contained is 75.0. This is illustrated graphically in Fig. 60, which shows the step from 75.0 to 79.9 with the ten frequencies contained within it numbered from 39 to 48 (inc.), together with the five steps in the step interval by which the median is measured. This shows that the point which constitutes the end of the thirty-ninth score and the beginning of the fortieth score lies exactly on the point which is 75.5.

The calculations for the median are shown in Table VIII. The first column gives the labels of the score groups in the distribution, the second gives the frequencies, and the third gives the successive additions to reach the step in which the

TABLE VIII. CALCULATIONS FOR FINDING THE MEDIAN

SCORE GROUPS	FREQUENCIES	ADDITION OF FREQUENCIES
110.0–114.9	1	
105.0–109.9	3	
100.0–104.9	5	
95.0– 99.9	4	
90.0– 94.9	7	
85.0– 89.9	4	
80.0– 84.9	6	
75.0– 79.9	10 Median lies in this step	
70.0– 74.9	10	38
65.0– 69.9	6	28
60.0– 64.9	5	22
55.0– 59.9	7	17
50.0– 54.9	5	10
45.0– 49.9	2	5
40.0– 44.9	2	3
35.0– 39.9	1	1
	78 $\left(\frac{N}{2} = 39\right)$	

Remainder = 1; frequencies in step in which median lies = 10; size of step interval = 5; beginning of step in which median lies = 75.0.

$\frac{1}{10} \times 5 + 75.0 = 75.50$. Median = 75.50.

median lies. At the foot of the table will be found calculations for determining the median.

The first quartile point. The first quartile point, usually called Q_1, is that point which separates the first, or lowest, quarter of a distribution from the second, or next higher, quarter. It is found in much the same way as is the median, with certain important changes in the early steps.

STEP 1. *Making a frequency distribution*

As in all the calculations for point measures, it is necessary to start by making a frequency distribution.

STEP 2. *Determining one quarter of the frequencies* $\left(\frac{N}{4}\right)$

One quarter of the frequencies of the distribution is found by adding the total number of frequencies and dividing the

sum by 4. In the example given on page 156 the total number of frequencies is 78. This divided by 4 gives a result of 19.5.

STEP 3. *Adding the frequencies up to the step in which Q_1 is contained*

Beginning with the frequency in the lowest group in the frequency distribution, the third step is to add successively the frequencies, group by group, up to that in which the first quartile point is contained.

In the case cited the frequencies total 1 for the lowest score group from 35.0 to 39.9, then 3 $(1 + 2)$, then 5 $(1 + 2 + 2)$,

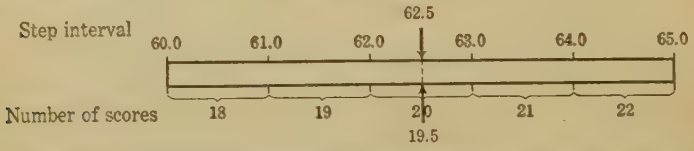


FIG. 61. Graphic representation of placement of first quartile point

then 10 $(1 + 2 + 2 + 5)$, and then 17 $(1 + 2 + 2 + 5 + 7)$. This leaves a remainder of 2.5; and since the next frequency in the score group from 60.0 to 64.9 is 5, Q_1 lies in that step.

STEP 4. *Determining Q_1*

Q_1 is found by a formula exactly like that used for finding the median. The remainder necessary to reach the quartile point determined by $\frac{N}{4}$ is divided by the number of frequencies in the step in which Q_1 lies, this is multiplied by the size of the step interval, and the result is added to the beginning of that step interval.

The remainder has been found to be 2.5 $(19.5 - 17)$, the number of frequencies in the step is 5, the size of the step interval is 5, and the beginning point of the step is 60.0.

$$\frac{2.5}{5} \times 5 + 60.0 = 62.5.$$

$$Q_1 = 62.5.$$

Fig. 61 shows the placement of the first quartile point in the step from 60.0 to 64.9.

Table IX shows the calculations necessary in finding the first quartile point. In the first column will be found the group labels for the frequency distribution, in the second column will be found the frequencies of that distribution, in the third column are the additions of the frequencies up to the step in which the first quartile point is contained, and at the bottom of the table are the necessary calculations for determining the quartile point desired.

TABLE IX. CALCULATIONS FOR FINDING THE FIRST QUARTILE POINT

SCORE GROUPS	FREQUENCIES	ADDITION OF FREQUENCIES
110.0–114.9	1	
105.0–109.9	3	
100.0–104.9	5	
95.0– 99.9	4	
90.0– 94.9	7	
85.0– 89.9	4	
80.0– 84.9	6	
75.0– 79.9	10	
70.0– 74.9	10	
65.0– 69.9	6	
60.0– 64.9	5 Quartile 1 lies in this step	
55.0– 59.9	7	17 (Remainder = 2.5)
50.0– 54.9	5	10
45.0– 49.9	2	5
40.0– 44.9	2	3
35.0– 39.9	1	1
	78	

Remainder = 2.5; frequency of step containing $Q_1 = 5$; size of step interval = 5; beginning point of step in which Q_1 is contained = 60.0.

$$\frac{2.5}{5} \times 5 + 60.0 = 62.5.$$

$$Q_1 = 62.5.$$

The third quartile point. The third quartile point, usually called Q_3, is that point which separates the third quarter of the distribution from the fourth, or highest, quarter. It

is found in the same way as the median or the Q_1, with changes only in the early steps.

STEP 1. *Constructing a frequency distribution*

This is a necessary preliminary to any calculation of point measures.

STEP 2. *Finding three quarters of the frequencies* $\left(\frac{3N}{4}\right)$

Three quarters of the frequencies is found by dividing the total number of frequencies by 4 and multiplying the result by 3. In the case cited above the total number of cases is 78, which gives 58.5 as the value of $\frac{3N}{4}$

STEP 3. *Adding the frequencies up to the step in which* Q_3 *is contained*

Beginning with the frequencies in the lowest score group and adding successively up the column to the group in which the third quartile point is contained shows that Q_3 is contained in the group 90.0–94.9, with a total of 58 to that point and a remainder of 0.5.

STEP 4. *Calculating the third quartile point* (Q_3)

The calculation for the third quartile point is similar to the two calculations already described. To the remainder of the frequencies necessary to reach $\frac{3N}{4}$, divided by the frequency of the step in which it is contained and multiplied by the size of the step interval, add the size of the beginning point of that step. The remainder has been found to be 0.5, the frequency of the step in which Q_3 is contained is 7, the size of the step interval is 5, and the beginning point of the step in which Q_3 is contained is 90.0. The calculations are as follows:

$$\frac{0.5}{7} \times 5 + 90.0 = 90.35.$$

$$Q_3 = 90.35.$$

Fig. 62 shows the placement of the third quartile in the step from 90.0 to 94.9.

Table X shows the calculations necessary for finding the third quartile point. The first column shows the labels of the

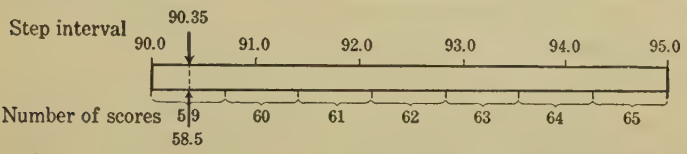


FIG. 62. Graphic representation of placement of third quartile point

score groups, the second the frequency for each score group, the third the successive additions up to the step in which Q_3 is contained, and at the bottom of the table are shown the necessary calculations for finding the third quartile point.

TABLE X. CALCULATIONS FOR FINDING THE THIRD QUARTILE POINT

SCORE GROUPS	FREQUENCIES	SUM OF FREQUENCIES
110.0–114.9	1	
105.0–109.9	3	
100.0–104.9	5	
95.0– 99.9	4	
90.0– 94.9	7 Quartile 3 lies in this step	
85.9– 89.9	4	
80.0– 84.9	6	
75.0– 79.9	10	
70.0– 74.9	10	
65.0– 69.9	6	
60.0– 64.9	5	
55.0– 59.9	7	
50.0– 54.9	5	
45.0– 49.9	2	
40.0– 44.9	2	
35.0– 39.9	1	
	78	

Remainder = 5; frequency in step in which Q_3 is contained = 7; size of step interval = 5; beginning point of step = 90.0.

$$\frac{0.5}{7} \times 5 + 90.0 = 90.35. \quad Q_3 = 90.35.$$

Special considerations in finding medians and quartiles. In this method of finding medians and quartiles there is but one situation which should give any particular difficulty and for which one should watch carefully. This occurs only when the remainder is zero, and when at the same time the number of frequencies in the step where the measure would usually lie is also zero. In this case the following procedure should be followed: Add to the beginning of the step interval half the amount of the step or steps to the next numbered frequency. Thus, if there is one zero frequency and the step interval is 5, the amount to be added to the beginning of the step in which the median would usually lie is 2.5. If there are two zero frequencies in successive score groups at this point and the step interval is 4, there would be an amount half the size of the entire amount to reach the next frequency (8), to add to the beginning of the step interval in which the median would usually lie; this amount would be 4. In the example given on page 179 a frequency distribution is shown in which the remainder in finding the amount of $\frac{N}{2}$ is 0 and the next step is also 0. The size of the step interval is 7, half of which is 3.5. This amount added to the beginning of the step in which the median would usually lie makes the calculation for the median as follows:

Remainder = 0.
Frequency in step usually containing the median = 0.
Step interval = 7.
Beginning point of step usually containing the median = 91.0.

$$91.0 + (\tfrac{1}{2} \times 7) = 94.5.$$

Median = 94.5.

The assumption in making this adjustment is simply that the scores of the last numbered frequency reach through the zero steps to locate the theoretical mid-point.

TABLE XI. CALCULATION OF MEDIAN IN FREQUENCY DISTRIBUTION CONTAINING ZERO FREQUENCY IN STEP USUALLY CONTAINING THE MEDIAN

SCORE GROUPS	FREQUENCIES	SUM OF FREQUENCIES
133.0–139.9	1	
126.0–123.9	0	
119.0–125.9	2	
112.0–118.9	4	
105.0–111.9	8	
98.0–104.9	9	
91.0– 97.9	0 Step usually containing median	
84.0– 90.9	10	24 (Remainder = 0)
77.0– 83.9	5	14
70.0– 76.9	3	9
63.0– 69.9	3	6
56.0– 62.9	0	3
49.0– 55.9	2	3
42.0– 48.9	1	1
	48	

$N = 48$; $\frac{N}{2} = 24$; remainder = 0; frequency in step usually containing the median = 0; step interval = 7; beginning point of step usually containing median = 91.0.

$$91.0 + (\tfrac{1}{2} \times 7) = 94.5.$$

Median = 94.5.

If there had been three consecutive score groups with frequencies of zero, the calculations would be as follows:

Remainder = 0.
Frequency in step usually containing the median = 0.
Frequencies in next two steps = 0.
Step interval = 7.
3 × step interval = 21 = total amount of zero interval.
Beginning point of step in which median would usually lie = 91.0.

$$91.0 + (\tfrac{1}{2} \times 21) = 101.5.$$

Median = 101.5.

At the foot of Table XII it will be noted that calculations are given for finding the first quartile point in the distribu-

TABLE XII. CALCULATIONS SHOWING DETERMINATION OF MEDIAN AND Q_1 IN A FREQUENCY DISTRIBUTION CONTAINING REMAINDERS OF ZERO AND WITH ZERO FREQUENCIES IN STEPS USUALLY CONTAINING MEDIAN AND FIRST QUARTILE POINT

SCORE GROUPS	FREQUENCIES	SUM OF FREQUENCIES FOR MEDIAN	SUM OF FREQUENCIES FOR Q_1
133.0–139.9	2		
126.0–132.9	5		
119.0–125.9	7		
112.0–118.9	10		
105.0–111.9	0	Additional steps containing zero frequencies	
98.0–104.9	0		
91.0– 97.9	0	Step in which M usually lies	
84.0– 90.9	7	24 (Remainder = 0)	
77.0– 83.9	5	17	
70.0– 76.9	0	12 Step usually containing Q_1	
63.0– 69.9	5	12	12 (Remainder = 0)
56.0– 62.9	4	7	7
49.0– 55.9	2	3	3
42.0– 48.9	1	1	1
	48		

For median: $N = 48$; $\frac{N}{2} = 24$; remainder $= 0$; frequency in step usually containing median $= 0$; number of score groups adjacent in addition $= 2$; size of step interval $= 7$; $3 \times 7 = 21$; beginning of step usually containing the median $= 91.0$.

$$91.0 + (\tfrac{1}{2} \times 21) = 101.5.$$

Median $= 101.5$.

For Q_1: $N = 48$; $\frac{N}{4} = 12$; remainder $= 0$; frequency in step usually containing $Q_1 = 0$; size of step interval $= 7$; beginning point of step usually containing $Q_1 = 70.0$.

$$70.0 + (\tfrac{1}{2} \times 7) = 73.5.$$

$Q_1 = 73.5$.

tion there given. It will also be noted that this first quartile point lies in a step that has a zero frequency and where there has been a remainder of zero. The same procedure as was followed in finding the median under similar conditions is followed here. When there is no remainder and the step in which the quartile point would usually lie has a zero frequency, it is

necessary to add to the beginning point of the step interval an amount equal to half the step interval between the beginning point of that step and the beginning point of the step containing the next numbered frequencies. If the case is as shown in Table XII, then $\frac{N}{4} = 12$, and in adding up the frequencies it can be seen that there is no remainder and that the next step is one of zero frequency. In this case the step interval to the beginning point of the next step which contains a numbered frequency is 7. This divided by 2 gives 3.5, which should be added to the beginning point of the score group containing Q_1. The calculations are as follows:

Remainder = 0.
Frequency in step usually containing $Q_1 = 0$.
Size of step interval = 7.
Beginning point of step usually containing $Q_1 = 70.0$.

$$70.0 + (\tfrac{1}{2} \times 7) = 73.5.$$
$$Q_1 = 73.5.$$

The mean. The mean is a measure which is somewhat more complex to find than the median, but is generally better understood. It is the familiar arithmetic average; but it is calculated from the mid-point of the step and not, as is the usual arithmetic mean, from the beginning of the step.

The mean may be calculated in two ways. When the number of cases is small, it may be found by adding together the scores and dividing their sum by the number of scores. The formula, similar to the arithmetic formula, is

$$\frac{\text{Sum of scores}}{\text{Number of scores}} = \text{mean.}$$

It should be remembered that the mean is calculated from the mid-points of the scores and not from the beginning points of the scores as was the case when the medians or quartile points were calculated.

Steps in finding the mean by the guessed-mean method. When the number of cases is large, it is usually more convenient and far less laborious to calculate the mean by the guessed-mean method.

STEP 1. *Making a frequency distribution*

As in finding the quartile points and the median, the first need in finding the mean is to make a frequency distribution of the scores for which it is to be found. This is done according to the method described on pages 150 ff.

STEP 2. *Estimating the location of the mean*

The second step is to make an estimate of the location of the mean. This can be done anywhere in the distribution, always at the mid-point of the step in which it is selected, and it is usually better to guess this estimate somewhere near the point where the true mean is likely to lie. In the example given on page 183, the mean has been guessed as being in the score group from 75.0 to 79.9. The mid-point of the step is 77.5, which therefore becomes the guessed mean ($\underline{M}$).

STEP 3. *Finding plus and minus step deviations from the guessed mean*

Working on the frequency-distribution table, the next step is to find the plus and minus step deviations from this guessed mean. A step deviation is simply the number of steps any score group lies from the guessed mean. A plus step deviation is one which lies at a point greater than the step of the guessed mean; a minus step deviation is one which lies at a point less. In practice this operation can be made quite mechanical, and the more mechanical it is made the more accurate it is likely to be.

In the third column, which can be headed d (for deviation), lines should be drawn above and below the step in which is the guessed mean. Starting from that step, the step above

TABLE XIII. MEAN (STEPS 1 AND 2)

Frequency Distribution showing the Position of Step containing the Guessed Mean

SCORE GROUPS	FREQUENCIES	
110.0–114.9	1	
105.0–109.9	3	
100.0–104.9	5	
95.0– 99.9	4	
90.0– 94.9	7	
85.0– 89.9	4	
80.0– 84.9	6	
75.0– 79.9	10	Guessed mean = 77.6
70.0– 74.9	10	
65.0– 69.9	6	
60.0– 64.9	5	
55.0– 59 9	7	
50.0– 54.9	5	
45.0– 49.9	2	
40.0– 44.9	2	
35.0– 39.9	1	
	78	

deviates one step. This makes the deviation for the score group $80.0 + 1$. The next step above deviates two steps from the step including the guessed mean, which makes that step have a step deviation of $+ 2$. In the same way all the steps above these should be numbered according to the size of their deviations from the step 75.0–79.9.

When this has been completed, the minus deviations should be filled in in the same way. The step 70.0–74.9 deviates one step minus, and its deviation is therefore $- 1$; the next step below on the table deviates two steps minus, and its deviation therefore becomes $- 2$. The remainder of the steps should be filled in in the same way.

It should be remembered that all steps, whether they have zero frequencies or not, should always be included, since *steps* not frequencies are being dealt with.

TABLE XIV. MEAN (STEP 3)

Plus and Minus Step Deviations from the Guessed Mean

SCORE GROUPS	FREQUENCIES	DEVIATIONS
110.0–114.9	1	+7
105.0–109.9	3	+6
100.0–104.9	5	+5
95.0– 99.9	4	+4
90.0– 94.9	7	+3
85.0– 89.9	4	+2
80.0– 84.9	6	+1
75.0– 79.9	10	
70.0– 74.9	10	−1
65.0– 69.9	6	−2
60.0– 64.9	5	−3
55.0– 59.9	7	−4
50.0– 54.9	5	−5
45.0– 49.9	2	−6
40.0– 44.9	2	−7
35.0– 39.9	1	−8
	78	

STEP 4. *Multiplying the frequencies by their deviations from the guessed mean*

The next step is to multiply the frequencies in each score group by the size of the deviation, regarding signs in each case. These multiplications can be placed most conveniently in a fourth column in the distribution, which can be labeled *fd* to indicate its character.

In the example given in Table XV the frequency in the step 80.0 — 84.9 is multiplied by + 1, giving an *fd* of 6; in the next group above the frequency 4 is multiplied by the deviation + 2, giving an *fd* of 8, and so on for the remaining plus deviations and frequencies.

The frequency in the step immediately below that containing the guessed mean is 10; multiplying this by — 1, the step deviation, gives an *fd* of — 10. The next score group below that has a frequency of 6, which, multiplied by the

step deviation -2, gives an fd of -12. This process is continued for all the minus-step deviations and frequencies.

STEP 5. *Finding the sums of the plus and minus fd columns*

The next operation is to find separately, regarding signs, the sums of the plus and minus columns of fd. This gives a total degree of frequency deviation from the step containing the guessed mean.

As shown in the space in Table XV, the sum of the plus deviations is 101. The sum of the minus deviations is -124, which can be written in at the foot of the fd column.

TABLE XV. MEAN (STEPS 4 AND 5)

Multiplication of Frequencies by their Deviations and the Sums of the Plus and Minus Columns of fd

SCORE GROUPS	f	d	fd
110.0–114.9	1	7	7
105.0–109.9	3	6	18
100.0–104.9	5	5	25
95.0– 99.9	4	4	16
90.0– 94.9	7	3	21
85.0– 89.9	4	2	8
80.0– 84.9	6	1	6
75.0– 79.9	10	Sum of $fd+ = 101$	
70.0– 74.9	10	−1	−10
65.0– 69.9	6	−2	−12
60.0– 64.9	5	−3	−15
55.0– 59.9	7	−4	−28
50.0– 54.9	5	−5	−25
45.0– 49.9	2	−6	−12
40.0– 44.9	2	−7	−14
35.0– 39.9	1	−8	−8
	78	Sum of $fd- = -124$	

STEP 6. *Determining the correction in order to find the true mean*

The correction which is desired as a result of making this step is one which when added to or subtracted from the guessed mean will result in the true mean.

The formula for finding this correction is as follows:

$$\left(\frac{(\text{Sum of } fd+) + (\text{sum of } fd-)}{\text{Number of frequencies}}\right) \times \text{size of step interval} = \text{correction.}$$

All the data necessary for finding the correction has been obtained in the foregoing steps. The sum of $fd+$ is 101; the sum of the $fd-$ is -124, the number of frequencies is 78, and the size of the step interval is 5. Substituting in the foregoing formula, the calculations are as follows:

$$\frac{101 + (-124)}{78} \times 5 = \frac{-23}{78} \times 5 = -1.45+.$$

The correction for the guessed mean is therefore $-1.45+$.

STEP 7. *Finding the true mean*

The true mean is found by adding to or subtracting from the guessed mean the correction which has been found. The formula ($\underline{M}$ = guessed mean; C = correction; M = true mean) is

$$\underline{M} + C = M.$$

In the case cited above and in Table XV, where the guessed mean is 77.5 and the correction is -1.45, the substitution is as follows:

$$77.5 + (-1.45) = 76.05.$$

It makes little difference where the guessed mean is located. In the cases shown in Table XVI, where succeeding columns show different points of guessing, the true mean is practically the same in all cases; where it is not the same the variation will be found to be in the second decimal place and is due to the error of not carrying the decimal places out farther. In ordinary practice it will be found sufficiently accurate to carry the decimal places out two places. The calculations for the four other examples given to show that the mean can be guessed in any score group, are as follows:

1. *Mean guessed at 72.5.* Substitution as follows:

$$\frac{141 + (-86)}{78} \times 5 = +3.50,$$

$$72.5 + 3.5 = 76.00.$$

2. *Mean guessed at 67.5.* Substitution as follows:

$$\frac{191 + (-58)}{78} \times 5 = +8.50,$$

$$67.50 + 8.50 = 76.00.$$

3. *Mean guessed at 82.5.* Substitution as follows:

$$\frac{71 + (-172)}{78} \times 5 = -6.45,$$

$$82.50 + (-6.45) = 76.05.$$

4. *Mean guessed at 97.5.* Substitution as follows:

$$\frac{14 + (-349)}{78} \times 5 = -21.45,$$

$$97.50 + (-21.45) = 76.05.$$

The frequencies, deviations, and multiplication of deviations by frequencies are given in successive columns in Table XVI.

Use of measures of mass and measures of points. Measures of mass and measures of points are more or less fundamental to the interpretation of test scores and are so closely related to one another that a consideration of their uses is combined in this section.

As has been stated, the chief use of measures of mass is to describe the general characteristics of the mass of raw scores which results for any class from any test.

The *order distribution* has wide use as a means of placing scores from both standard and classroom tests in a form which enables one to transpose them into grades or marks. It is necessary also to make an order distribution as a preliminary to the making of a rank distribution. A third use

TABLE XVI. SHOWING CONSTANCY OF THE TRUE MEAN WITH THE GUESSED MEAN LOCATED IN DIFFERENT STEPS

SCORE GROUPS	*f*	$\underline{M} = 72.50$		$\underline{M} = 67.50$		$\underline{M} = 82.50$		$\underline{M} = 97.50$	
		d	*fd*	*d*	*fd*	*d*	*fd*	*d*	*fd*
110.0–114.9	1	8	8	9	9	6	6	3	3
105.0–109.9	3	7	21	8	24	5	15	2	6
100.0–104.9	5	6	30	7	35	4	20	1	5
95.0– 99.9	4	5	20	6	24	3	12	$fd - = 14$	
90.0– 94.9	7	4	28	5	35	2	14	−1	−7
85.0– 89.9	4	3	12	4	16	1	4	−2	−8
80.0– 84.9	6	2	12	3	18	$fd - = 71$		−3	−18
75.0– 79.9	10	1	10	2	20	−1	10	−4	−40
70.0– 74.9	10	$fd + = 141$		1	10	−2	20	−5	−50
65.0– 69.9	6	−1	−6	$fd + = 191$		−3	18	−6	−36
60.0– 64.9	5	−2	−10	−1	−5	−4	20	−7	−35
55.0– 59.9	7	−3	−21	−2	−14	−5	35	−8	−56
50.0– 54.9	5	−4	−20	−3	−15	−6	30	−9	−45
45.0– 49.9	2	−5	−10	−4	−8	−7	14	−10	−20
40.0– 44.9	2	−6	−12	−5	−10	−8	16	−11	−22
35.0– 39.9	1	−7	−7	−6	−6	−9	9	−12	−12
	78	$fd - = -86$		$fd - = -58$		$fd - = -172$		$fd - = -349$	
		$M = 76.00$		$M = 76.00$		$M = 76.05$		$M = 76.05$	

will be considered in a later chapter in connection with the discussion of problems in the classifying of pupils by means of standard tests.[1]

The *rank distribution* finds its chief use in the explanation of the true placement of the individual pupils in rank form within the range of scores of a test. Inasmuch as test ranges, even for the same group of pupils, change greatly in accordance with the forms of tests used, the rank distribution reduces that variation to a more intelligible ranking. Thus by mere inspection it is simple to see that if a child achieves a raw score of 35 in one test and 140 in another, and if his rank in each test is 20, his general achievement with respect to the achievement of the group of which he is a part is probably

[1] See Chapter XIII.

equivalent in the two tests and not, as might at first seem probable, as the difference between 35 and 140.

A second use for the rank distribution, particularly for those who use classroom tests, consists in using it as one of the steps in turning raw scores into M scores.[1]

The *frequency distribution* finds its greatest use in the fact that it is the necessary preliminary to the finding of all the point measures which have been considered in this chapter. Before it is possible to find medians, means, quartiles, or crude modes, it is necessary to make a frequency distribution.

The *frequency surface* finds its greatest use as a graphic measure for the interpretation of class achievement. This is a matter of careful judgment on the part of the person who wishes to use the frequency surface for this purpose, but the effort necessary to make a careful analysis of the surface which has been obtained for any group of pupils for any test will go far toward explaining some of the fundamental characteristics of that test as it respects the achievement of the pupils.

Although practically every frequency surface will be somewhat different from others, there are certain characteristics of these surfaces by which they can be classed and through which they can be interpreted. The types of curves that are likely to be found, together with their characteristics, follow.

Types of curves. 1. *The normal curve.* The true normal curve is a parabolic form of curve which rises gradually from the lower portion of the base line, reaches its height, and then falls gradually again to the base line. It never quite touches the base line, or rather it approaches the base line more closely the farther it is extended, and touches it at infinity. The normal curve is not actually found in school measurement; but when large numbers of individuals are measured in a trait which is complex or in which the causes that result

[1] Charles Russell, *Classroom Tests*, pp. 270–271.

in the trait are numerous, the resultant surface approaches closely to that of the normal curve. This is the case in the frequency surface of the measures of intelligence of school children, of measurement of physical traits such as height or weight, and, to a lesser extent, the measurement of school achievement.

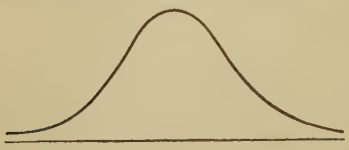


FIG. 63. The normal curve

The normal curve has the general appearance illustrated in Fig. 63. This is a "smoothed" curve. A smoothed curve is one where the irregularities of the various columns have been removed by drawing a line through the extremities which best show the characteristics of the entire surface. Such a curve is shown in Fig. 64, which represents the smoothing of the frequency surface as derived in Fig. 53.

FIG. 64. Smoothed curve of frequency surface, Fig. 53

2. *The plus skewed curve.* The plus skewed curve is one which rises rapidly from the lower portion of the base line to its height and then drops gradually to the upper portion of the base line, as in Fig. 65.

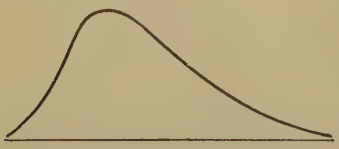


FIG. 65. A plus skewed curve

3. *The minus skewed curve.* The minus skewed curve is the opposite form to the plus skew. This curve rises gradually from the lower part of the base line, reaches its height, and then drops abruptly to the base line, as shown in Fig. 66.

4. *The multimodal curve.* This curve is very variable, and practically every curve which is found in practice will seem

to have some of its characteristics. In general the multimodal curve is one which seems to be made up of two or more distinct curves in combination. It may have somewhat the appearance of Fig. 67. It may also have more modes, and the modes may be more or less distinct.

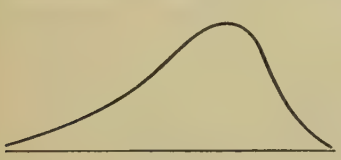


FIG. 66. A minus skewed curve

Interpretation of frequency surfaces. The interpretation of frequency surfaces resulting from standard tests differs from the interpretation of similar curves resulting from classroom tests.[1] In classroom tests the frequency surface is a measure of the fitness of a test to the pupils, or the measure of the relative difficulty of the test; in standard tests it is a measure of the more or less absolute achievement of the pupils. The extent of this measurement varies with the different standard tests in terms of the degree to which they are absolute measures. For this reason the interpretation of the frequency surface in standard tests must take into account the ability of the pupils, and the meaning of the frequency surface is dependent upon this ability. In Table XVII are some interpretations which can be made of various types of frequency surfaces. These interpretations must be modified in practice by the extent to which the actual curves differ from those here discussed.

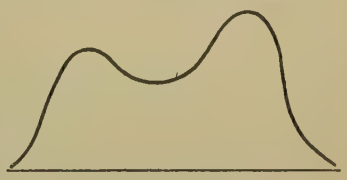


FIG. 67. A multimodal curve

When the frequency surface of ability is a curve that approximates a normal curve, there are three general conditions in the standard test curve which will influence any interpretation: (1) if the standard test curve is of normal type,

[1] Charles Russell, *Classroom Tests*, chap. xi.

(2) if it has a plus skew, or (3) if it has a minus skew. Within each of these type groups there are three further conditions which govern any interpretation; that is, whether the average, mean or median as may be preferred, is equal to, is less than, or is greater than the average measure of the ability curve. In Table XVII are shown the varieties of relations between ability and test curves, taking into account the averages. To use the table, an ability curve should first be made for any class of pupils. With that as a basis the median should be calculated and drawn on the curve, as is shown in Table XVII; then the actual test curve should be made and superimposed upon the ability curve, as in Table XVII.

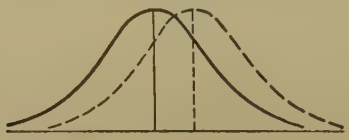
The data for the ability curve and the test curve must be in the same kind of units (not raw scores); otherwise it is impossible to make calculations of the mean, or median, which will be comparable for the two curves.

TABLE XVII. INTERPRETATION OF FREQUENCY SURFACES

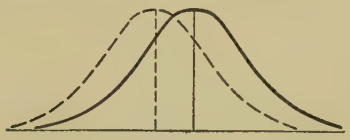
(Ability curves and medians in solid lines; achievement curves and medians in broken lines)

1. When the ability curve is normal
   a. And achievement curve is normal
      (1) And achievement median is same as ability median
      All pupils are working normally

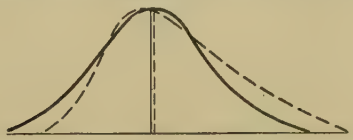
      (2) And achievement median is above ability median
      Entire group is above normal



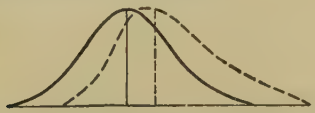
(3) And achievement median is below ability median
Entire group is below normal



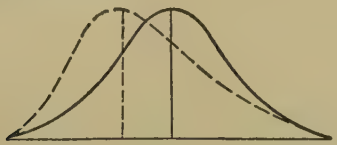
b. And the achievement curve is plus skew
(1) And achievement median is same as ability median
Best and poorest pupils are above normal; remainder normal



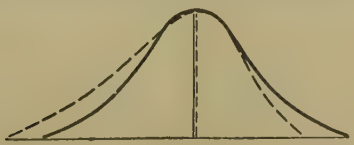
(2) And achievement median is above ability median
All pupils are above normal; best pupils farther above their normal than are poorer pupils



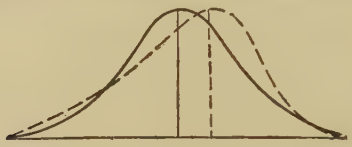
(3) And achievement median is below ability median
Almost entire group is below normal; poorer pupils farther below than remainder



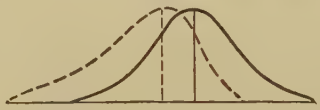
c. And the achievement curve is plus skew
(1) And achievement median is same as ability median
Best and poorest pupils are below normal; remainder normal



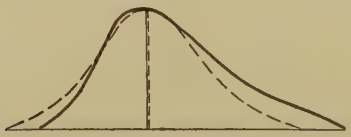
(2) And achievement median is above ability median
Average pupils are above their normal; remainder the same



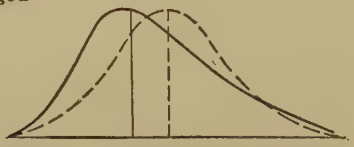
(3) And achievement median is below ability median
All pupils are below normal; poorer pupils farther below than best pupils



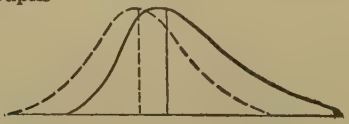
2. When the ability curve is plus skew
a. And the achievement curve is normal
(1) And the achievement median is same as ability median
Best and poorest pupils are below normal; remainder normal



(2) And the achievement median above ability median
Average pupils are above normal; best and poorest pupils unchanged

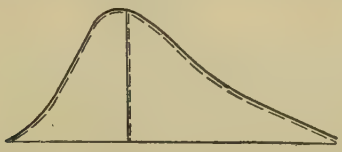


(3) And achievement median below ability median
All pupils are below normal; better pupils farther below than poorer pupils

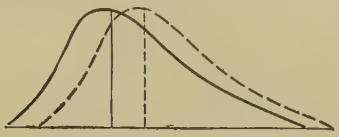


b. And the achievement curve is plus skew

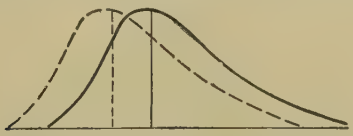
(1) And achievement median is same as ability median
Entire group is normal



(2) And achievement median is above ability median
Entire group is above normal

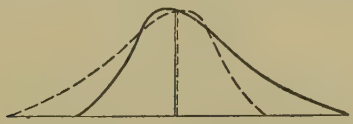


(3) And achievement median is below ability median
Entire group is below normal

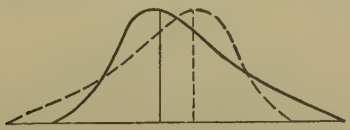


c. And the achievement curve is minus skew

(1) And achievement median is same as ability median
Average pupils are normal; best and poorest below normal



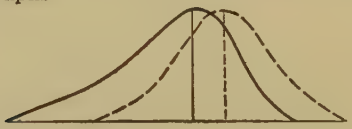
(2) And achievement median is above ability median
Average pupils are normal or above; poorest and best pupils below normal



(3) And achievement median is below ability median
All pupils are below normal; poorer pupils farther below than better pupils

3. When the ability curve is minus skew
a. And the achievement curve is normal
(1) And achievement median is same as ability median
Average pupils are normal; best and poorest pupils above normal

(2) And achievement median is above ability median
All pupils are above normal; better pupils farther above than poorer pupils

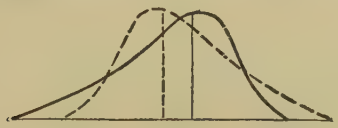


(3) And achievement median is below ability median
Average pupils are below normal; best and poorest pupils normal

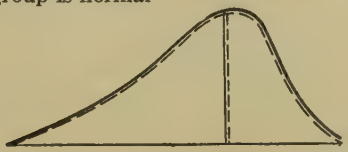
b. And the achievement curve is plus skew
(1) And achievement median is same as ability median
Average pupils are normal; best and poorest pupils above normal

(2) And achievement median is above ability median
All pupils are above normal; better pupils farther above than poorer pupils

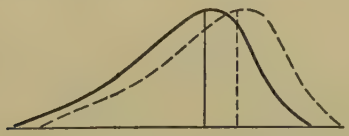
(3) And achievement median is below ability median
Average pupils are below normal; best and poorest pupils above normal



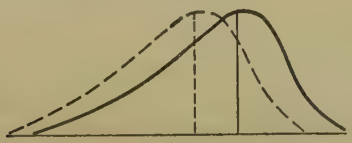
c. And the achievement curve is minus skew
(1) And achievement median is same as ability median
Entire group is normal



(2) And achievement median is above ability median
Entire group is above normal



(3) And achievement median is below ability median
Entire group is below normal



Interpretation of measures of points. Measures of points are much in general use as descriptive of the work of pupils, and the teacher should understand when to use them and how to determine their relative values.

As has been stated, the *crude mode* is an unreliable measure, but for the purpose of quick inspection it is frequently of value.

The *median* is somewhat more simple to find than the mean, and therefore has a somewhat more extended use. It is the measure to use when the true middle measure in any group of scores is desired. It is also the measure to use when it is desirable to know the *score of the average pupil* (compare with mean). When the norms of a test are given in terms of medians, as they often are, then the median should be used in order to have a measure that can be compared with those norms. When a curve is widely skewed either plus or minus, and the size of the extreme measures is of less concern than their number, then the median should be used in preference to the mean, since the latter is greatly influenced by the size of the extreme measures as well as by their placement.

The *mean* should be used whenever the norms of the test are given in terms of means. This is an indication that the size of the extreme scores should be taken into account, and that therefore the median is not an adequate measure. The mean should also be used whenever it is desirable to find the *average score of all the pupils* (compare with median). When the size of the extreme measures is as important to consider as is their number, then the mean should be used rather than the median.

Quartile points are measures of the range of the test scores and are frequently of value when the work of two groups of pupils is to be compared. The two groups may have practically the same medians or means, but one group may be much more compact than the other and therefore present

quite a different problem from the standpoint of teaching. Quartile points will reveal these differences.

Chapter summary. Measures of points describe the character of a group of scores in terms of what may be thought of as averages. The measures of points considered are the crude mode, the median, quartile points, and the mean.

The crude mode is simply the most frequent measure and may be found by an inspection of the scores in a frequency distribution.

The median is found by making a frequency distribution, determining half the frequencies, adding the frequencies up to the step in which the median is contained, and determining the median by dividing the remainder of the frequencies to reach half the total by the size of the frequency in the step containing the median. The whole is then multiplied by the size of the step interval and added to the beginning point of the step.

The quartile points are found by the same method, substituting a different formula than that used in finding the median. For Q_1, where $\frac{N}{2}$ is used for the median, $\frac{N}{4}$ is used; for Q_3, where $\frac{N}{2}$ is used for the median, $\frac{3N}{4}$ is used; in each case the remainder of the process and the remainder of the formula are the same.

The mean is found by locating a guessed mean at the mid-point of some step in a frequency distribution, making a table of plus and minus step deviations from that guessed mean, multiplying the frequencies in each step (regarding signs) by the size of the step deviations, and from that calculating a correction which added to, or subtracted from, the guessed mean gives the true mean.

Since measures of mass and measures of points are closely interrelated in the interpretation of the real meaning of test scores, the use and meaning of each measure of mass and measure of points is discussed in closing this chapter.

SELECTED BIBLIOGRAPHY

FENTON, N., and WORCESTER, D. A. An Introduction to Educational Measurements, pp. 109–112. Ginn and Company, Boston, 1928.

GREENE, H. A., and JORGENSEN, A. N. The Use and Interpretation of Educational Tests, chap. vi. Longmans, Green & Co., New York, 1929.

GREGORY, C. A. Fundamentals of Educational Measurement, chap. x. D. Appleton and Company, New York, 1922.

MCCALL, W. A. How to Measure in Education, chap. xv. The Macmillan Company, New York, 1922.

OTIS, A. S. Statistical Method in Educational Measurement, chap. ii. World Book Company, Yonkers-on-Hudson, 1925.

RUGG, H. O. Statistical Methods applied to Education, chap. v. Houghton Mifflin Company, Boston, 1917.

TRABUE, M. R. Measuring Results in Education, chap. ix. American Book Company, New York, 1924.

CHAPTER IX

MEASURES OF ACHIEVEMENT

Meaning of a scale. Whenever there are quantities to be measured, it is necessary to find some standard unit by which to express the measurement. For example, a stick may have a certain length; but if one merely says it is the length of a stick, that means nothing. Similarly, a rope may be said to have a certain length; but if that is expressed merely as "the length of a rope," no one can tell how long the rope is. In linear measure many units are used; probably the most familiar are those of inches, feet, and yards. In these terms the stick is measured as four feet long and the rope as ten yards long. Thus the length of a yard being known, there is an understanding of the length of the stick or the rope.

In standard tests it is necessary, as has already been stated, to transpose the raw scores of tests into other and more intelligible scores. To do this it is necessary to have some standard scale of measurement, or some standard unit on a standard scale of measurement, into which the scores can be changed, just as it is necessary to have some unit of length by which to compare the lengths of the rope and the stick.

As in linear measurement there are several different units of length, so in educational measurement there are several units or standards by which the products of pupils may be measured. These units, selected because of certain qualities which they possess in rendering raw scores meaningful, are called scales. They are the units analogous to feet, inches, and yards, with which a teacher should be familiar.

Types of scales. There are several types of scales in addition to the subject-matter scales which are now in general use among teachers to describe the work of pupils. The norms, which were discussed in Chapter V, are published in terms of one or more of these scales.

One scale which is very generally understood by teachers is called the *grade scale*. In this scale the units used are grades and fractions of grades. Thus the norm may be given in the form "A raw score of 32 means a grade level of 4.2." Translated into words, this statement means that a child who achieves a score of 32 points in this given test has accomplished as much as the average child at the end of the second month of fourth grade.

A second scale in general use is called the *age scale*. In this scale the standard set up is the achievement of pupils of any given age. Thus, if in a certain test the norm is given in terms of the age scale as "A score of 47 is equivalent to an age of 12–7," it means that an achievement of 47 points in the test is equivalent to the achievement of the average pupil who is twelve years and seven months old.

A third scale which is used is called the *T scale*. This scale has certain advantages not possessed by either the grade scale or the age scale, which will be discussed in later sections where the derivation and use of the scale are explained. The T score is the average score of twelve-year-old children, and therefore the standard is the achievement of twelve-year-old children. A raw score transposed into a T score shows how far above or below the average of twelve-year-old pupils is the score in question. There are other scales similar to the T scale based on dispersion of a group, such as the C scale of Van Wagenen and Pintner's scale of mental indexes. In high schools the Otis percentile norms are much used.

MEASURES OF ACHIEVEMENT

Measures of achievement in the grade scale. The unit in the grade scale is the *grade score,* usually referred to as the G score. In this scale, to say that a pupil has a G score in reading of 4.5 is to say that in reading he has an achievement equivalent to the average achievement of pupils in reading at the end of the fifth month of fourth grade. If, then, the raw scores of pupils are transposed into the G scores for the tests that have been given, a set of scores will result that will show the grade levels of the pupils in the different subjects.

Transmutation of raw scores into G scores. When the raw scores of tests are to be changed into G scores the following procedure should be used.

STEP 1. *Determining the grade norms for the test*

As already stated, practically every test has a list of grade or other norms published with it either in the manual of directions or in a separate sheet. The first step in transposing raw scores into G scores, therefore, is to determine the grade norms as given in the manual of the test. The grade norms for a test were given as shown in Table XVIII.

TABLE XVIII. GRADE NORMS FOR TEST A

January

GRADE	RAW SCORE
III	10
IV	25
V	36
VI	40
VII	49
VIII	56

It will be noted that the norms are the achievements of pupils in the third grade through the eighth in the month of January. This means that in the month of January the children in the third grade in the schools from which the norms were derived achieved an average score of 10, that the children in the fourth grade achieved an average score of 25, and so on.

STEP 2. *Preparing a transmutation sheet*

With the given norms as a basis, a table of G scores corresponding to every raw score made by any pupil can be made. To do this it is necessary to make a transmutation sheet to determine the various scores equivalent to the various months in each grade, as is shown in the following pages.

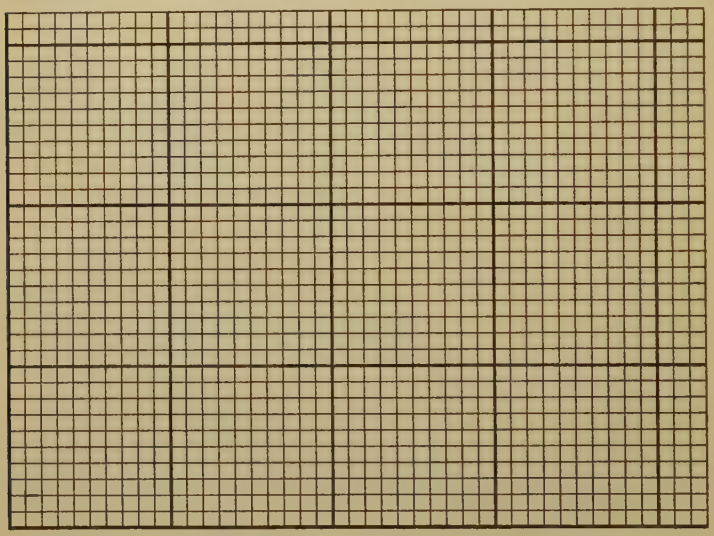


FIG. 68. Blank transmutation sheet

The transmutation sheet is one of ordinary graph paper, such as was used in the making of the frequency surface. It is wise to use for this purpose a printed graph paper, rather than to attempt to rule the paper oneself, since the accuracy of the transmutation depends on the accuracy of the lines of the graph paper. It is also a help to use a graph sheet which is ruled in squares of ten units, as this makes the reading of the graph simpler. Fig. 68 shows the appearance of the graph paper before preparing it for transmutation of raw scores.

The sheet being laid with the long side horizontal, the columns should be labeled horizontally, beginning at the extreme left-hand lower corner, to correspond with the grade units. Thus the first heavy-line ten columns to the right would be labeled "1.0," the next heavy line to the right would be labeled "2.0," and so the intermediate lines, starting with 0, would be 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, etc. In practice it is necessary to place the beginning point of each grade only, as the intermediate steps can easily be followed and understood.

With this completed the sheet should now be labeled, beginning at the lower left-hand corner and running vertically up the left-hand side of the sheet, with the various raw scores of the test, allowing one space for each raw-score unit. Thus the first heavy line above the 0 would be labeled "10," the next above would be labeled "20," and so on, to the extent of the raw scores. The intermediate lines should either be labeled or understood to indicate 1, 2, 3, 4, etc., 11, 12, 13, 14, etc., and so on.

When the transmutation sheet has been completely prepared, it should have the appearance shown in Fig. 69.

STEP 3. *Placing the norm positions on the transmutation sheet*

It is now possible to place the norm points on the transmutation sheet. Beginning with the norm for Grade III and referring to the norms as given in Table XVIII, the norm is 10. This is for the month of January. In this case there is no statement whether the norm is for the beginning or the end of January, or for the entire month, and it is therefore necessary to make an assumption. It is safest to assume in this case that the norms as given represent the average scores of pupils in the third grade throughout the month of January, and that therefore the best point which can be used to represent this average is the halfway point between the first of January and

the first of February. Assuming also that the schools from which these norms were derived began in September, the middle of January represents a period of instruction of four and one-half months beyond the beginning of third grade for the pupils from whose achievement these norms were obtained.

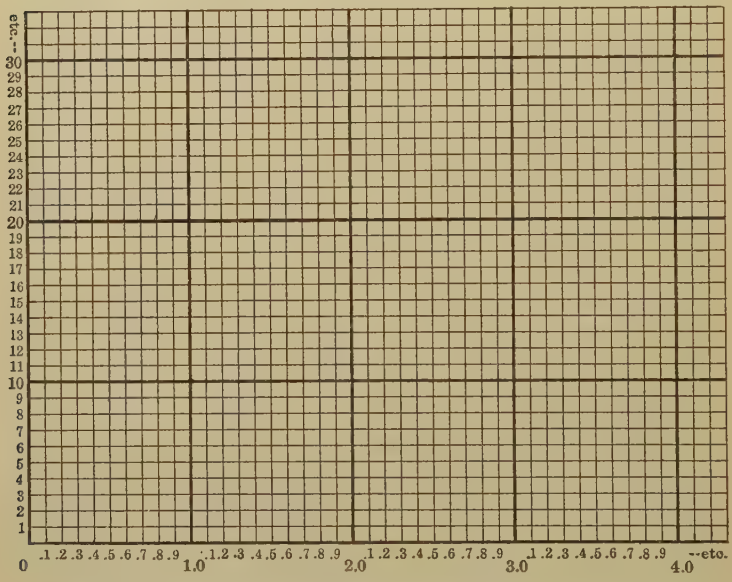


FIG. 69. Portion of transmutation sheet (labeling of columns horizontally and vertically)

By this process of reasoning, which should be followed whenever raw scores are to be transmuted into G scores, it is clear that a score of 10 really means the average achievement of pupils who have completed four and one-half months of the third grade, or a G score of 3.45.

This point should be placed on the transmutation sheet. Starting on the left-hand vertical column, which represents the raw scores, the line indicating 10 should be located. This

line should be followed across horizontally until a point is reached directly above the line labeled, or understood as being halfway between 3.4 and 3.5. This point is labeled "A" in Fig. 70.

The grade norm for fourth grade is given in Table XVIII as 25. Following the same reasoning as in the case just described, and assuming that the norm represents the average

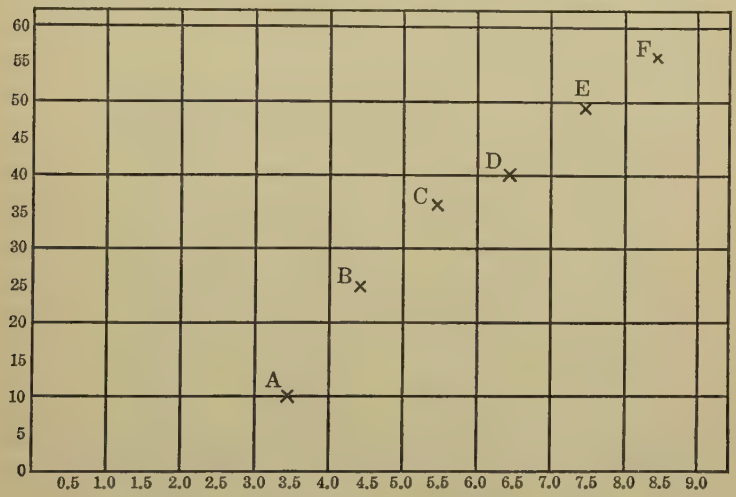


FIG. 70. Placement of norm positions on transmutation sheet

achievement of pupils in the fourth grade during January and that the best point to represent this is 4.45, it must also be assumed that an achievement of 25 is the average achievement of pupils at the end of four and one-half months of fourth grade. This point can now be placed on the transmutation sheet. Following up the left-hand margin until the point representing a raw score of 25 is reached, this line should be followed out horizontally until it passes over a point which represents a G score of 4.45, halfway between 4.4 and 4.5. This point is labeled "B" in Fig. 70.

In the same manner all the remaining grade norms should be placed on the transmutation sheet: that for the fifth grade at a point where the lines 36 and 5.45 cross, labeled "C" in Fig. 70; that for the sixth grade where the lines of 40 and 6.45 cross, labeled "D"; that for the seventh grade where the lines of 49 and 7.45 cross, labeled "E"; and, finally, that for the eighth grade where the lines of 56 and 8.45 cross, labeled "F."

The appearance of the transmutation sheet at this point is given in Fig. 70.

STEP 4. *Drawing the norm line of the transmutation sheet*

The next step is to draw the norm line on the transmutation sheet. This is done by drawing a smoothed line (one

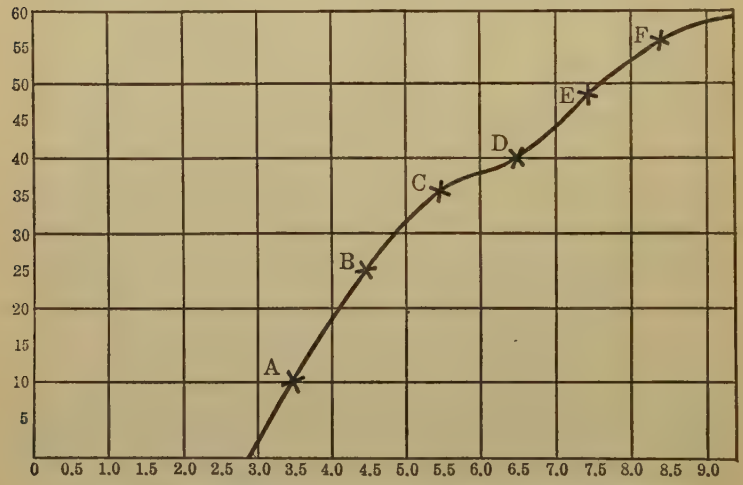


FIG. 71. Drawing of norm line on transmutation sheet

that has no sharp angles) through the norm points that have been determined. This line shows the average development of achievement in the test throughout all the grades, and shows that this development is not uniform throughout but

changes at times, proceeding rather rapidly in the third and fourth grades, becoming less rapid in the fifth and sixth grades, and accelerated again in the seventh and eighth grades. The norm line should be produced from point *A* in the same general direction as from *B* to *A* until the norm line crosses the zero point of the test. For this norm line this occurs at 2.9, which shows that this particular test is unfitted for pupils below that point, nearly the beginning of the third grade. The norm line should also be extended beyond point *F* in the same direction as it takes between points *E* and *F*, until it crosses the point at the end of the eighth grade, 8.9. If necessary it can be produced somewhat beyond that point at the same angle if certain ninth-grade norms are desired.

The norm line for this test, when drawn as indicated, has the appearance shown in Fig. 71. The points previously indicated are shown with crosses and are lettered.

STEP 5. *Developing a raw-score-G table from the transmutation sheet*

It is now possible to make up a table of G scores for all raw scores on the original test papers. An easy way to do this is to make a tabulation of all the possible raw scores from 1 to the upper limit of the grade scores — in this case 59 for grade 9.2.

Beginning at the left-hand side and following the line of 1 across horizontally it will be seen that the line crosses the norm line at a point approximately 2.9 on the grade scale below. This can therefore be transposed to the table opposite the score of 1. Following the line of 2 across in the same way, it crosses the norm line at a point close to a grade score of 3.0, which can then be transferred to the new table. A score of 3 crosses approximately at 3.1, and a score of 4 crosses the norm line at almost exactly 3.1. In this case the scores of 3 and 4 should both be considered as having G scores of 3.1.

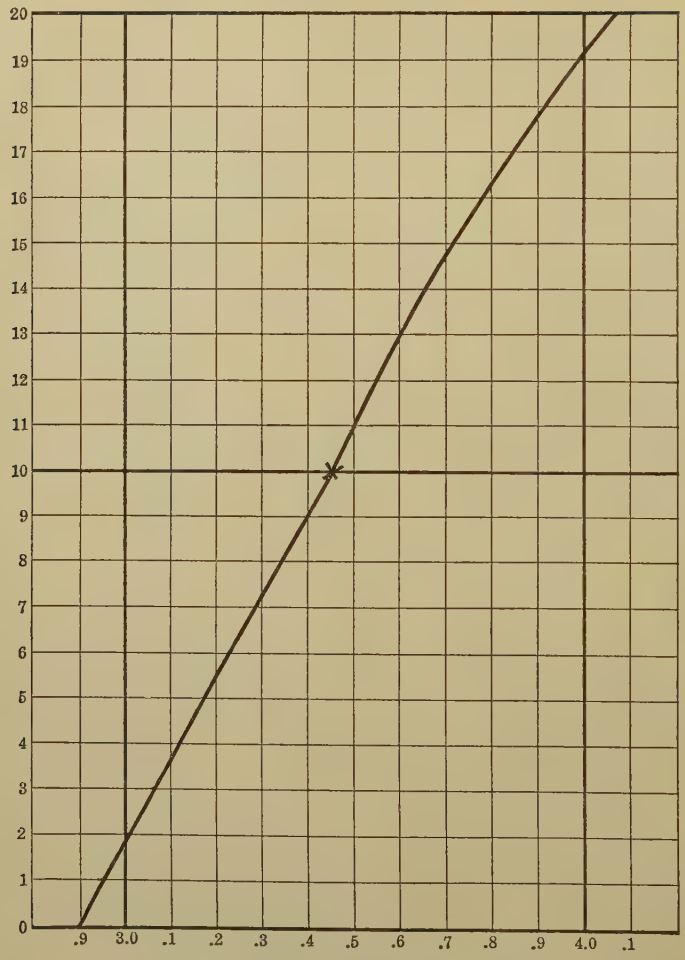


FIG. 72. Enlargement of portion of norm line

An enlargement of this part of the transmutation sheet will show how G scores are derived from raw scores of from 1 to 19. This enlargement is shown as Fig. 72, and a comparison with Fig. 71 will show the portion reproduced.

The process of locating the points where the horizontal line for each score crosses the G-score lines at the point of intersection with the norm line should be continued until the G-score table is constructed. This table when completed will have the appearance of Table XIX.

TABLE XIX. COMPLETED RAW-SCORE-G TABLE FOR TEST WITH NORMS AS GIVEN IN TABLE XVIII

SCORE	GRADE	SCORE	GRADE
1	2.9	31	4.9
2	3.0	32	5.0
3	3.1	33	5.1
4	3.1	34	5.2
5	3.2	35	5.3
6	3.2	36	5.5
7	3.3	37	5.6
8	3.3	38	5.8
9	3.4	39	6.2
10	3.5	40	6.5
11	3.5	41	6.6
12	3.6	42	6.7
13	3.6	43	6.8
14	3.7	44	6.9
15	3.7	45	7.0
16	3.8	46	7.1
17	3.8	47	7.2
18	3.9	48	7.3
19	4.0	49	7.5
20	4.1	50	7.6
21	4.1	51	7.7
22	4.2	52	7.8
23	4.3	53	8.0
24	4.4	54	8.1
25	4.5	55	8.3
26	4.5	56	8.5
27	4.6	57	8.6
28	4.7	58	8.9
29	4.8	59	9.2
30	4.8		

Achievement measures in the age scale. The unit of measurement in the age scale is an age *score*. This score is in terms of years and months and represents the achievement of pupils in terms of age. To say that a pupil has a reading age of 10–5 means that, regardless of his chronological age, he is able to read as well as the average pupil who is ten years and five months old. To determine this age as the result of a test, it is necessary to transmute the raw scores on the test into age units.

Transmutation of raw scores into age scores. To transmute any raw score on a test into an age score, it is necessary to make for the test a raw-score-age table which will show for any raw score that a pupil might achieve the equivalent age in the subject of that test.

STEP 1. *Listing the age norms*

In this transmutation it is assumed that a list of age norms is furnished with the test. These age norms are found in the direction booklet or sheet of directions for giving and interpreting the test, as was described in Chapter V. A later section will be devoted to the derivation of age scores from raw scores when the age norms are not given.

TABLE XX. LIST OF RAW-SCORE-AGE NORMS AS TAKEN FROM A DIRECTION SHEET OF A STANDARD TEST

RAW SCORE	AGE
8	7–0
18	8–0
25	9–0
28	10–0
39	11–0
48	12–0
52	13–0
59	14–0
67	15–0

The age norms should be listed, then, as they are given in the direction sheet. In the sample here used the age norms for the test were given in the direction sheet as listed in Table XX. This table can be read as follows: For a score of 8 the average age or age norm of achievement in this test

is that of pupils who are seven years and no months old. The purpose of the transmutation of the scores is to determine the ages for all the intervening scores that are not given in the table of norms; that is, to determine the age norm for a score of 9, 10, 11, and so on.

STEP 2. *Preparing a transmutation sheet*

As in the development of the raw-score-grade table described on page 203, it is necessary to prepare a transmutation

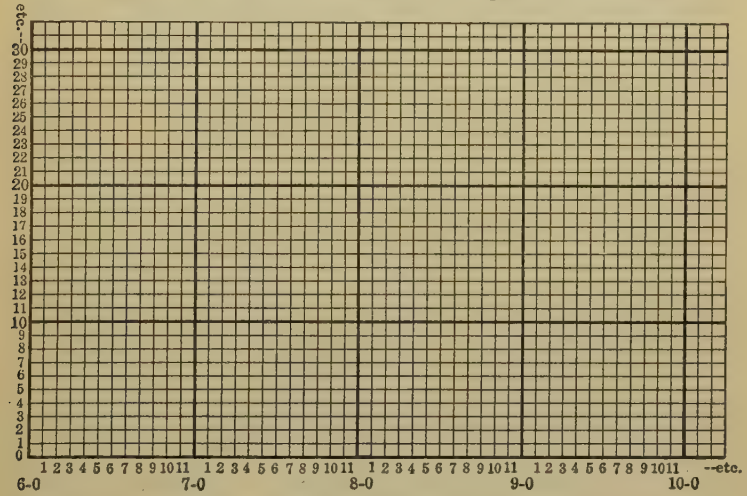


FIG. 73. Portion of transmutation sheet prepared for age norms

sheet. This is done by taking a sheet of cross-section paper similar to that described on page 204. The sheet being laid horizontally, the columns at the foot of the sheet should be labeled from left to right, starting with an age of six years and no months, which is somewhat less than a zero score in the test. Each column should indicate one month, and the yearly columns, 7–0, 8–0, 9–0, etc., can be more clearly noted if the vertical lines that indicate those months are emphasized as in Fig. 73.

When the horizontal columns have been labeled, the vertical columns should be labeled with the raw scores of the test. By referring to Table XX it will be seen that the scores on the test run from 8 to 67. Beginning, then, with the lower left-hand corner of the sheet, each vertical column should be labeled to indicate one score. It is not necessary to label every score if it is understood that every score is included. Thus, counting up ten spaces, the tenth should be labeled to indicate a score of 10, counting up ten more, the space should be labeled to indicate a score of 20, and so on until a score of 70 is reached.

The completed sheet ready for further steps will have the appearance of Fig. 73.

STEP 3. *Placing norm points on the transmutation sheet*

The norm points, as given in Table XX, can now be placed on the transmutation sheet. Following up the vertical column until the line is reached which is labeled or understood as "8," that line should be followed across until it crosses, at the foot of the sheet, the line labeled "7–0." In Fig. 74 this point is indicated by a small cross and the letter A is written in. The next norm point is where the line 18, in the left-hand column indicating raw scores, crosses the line 8–0 in the horizontal column indicating ages. This point can be crossed, as in Fig. 74, and the letter B written in to make it clear. Following this, all the points should be placed on the transmutation sheet — 25 where the line crosses the age line 9–0, 28 where the line crosses the age line 10–0, and so on.

When this step has been completed, the transmutation sheet will appear as in Fig. 74.

STEP 4. *Drawing the norm line*

The norm line can now be drawn. This is done similarly to the operation described on page 208, by drawing a smoothed curve through the points which have been placed

on the transmutation sheet. When this curve has been drawn, the ends should be extended from the point labeled "A" to a point crossing the line of zero scores. This extension should have the same general curve as that portion of the norm line between the points B and A. The other end of the line should be extended in the same general direction as

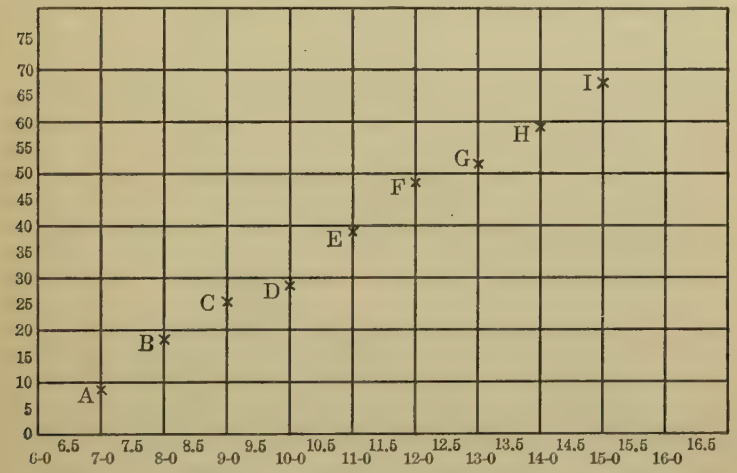


FIG. 74. Placement of norm points on transmutation sheet

the curve from H to I until it crosses the age-line indicated 16–0. This gives a completed norm line, as shown in Fig. 75.

STEP 5. *Constructing the raw-score-age table*

All preparations have now been made to construct the age table that is desired. A sheet of paper should be prepared containing in one column all the possible raw scores from 1 to 74. The score of zero should not be included, because when a score of zero is achieved one does not know whether the true score is really zero and equal therefore to an age which is a little less than the age for the score of 1, or whether the true score and its corresponding age are far less than zero.

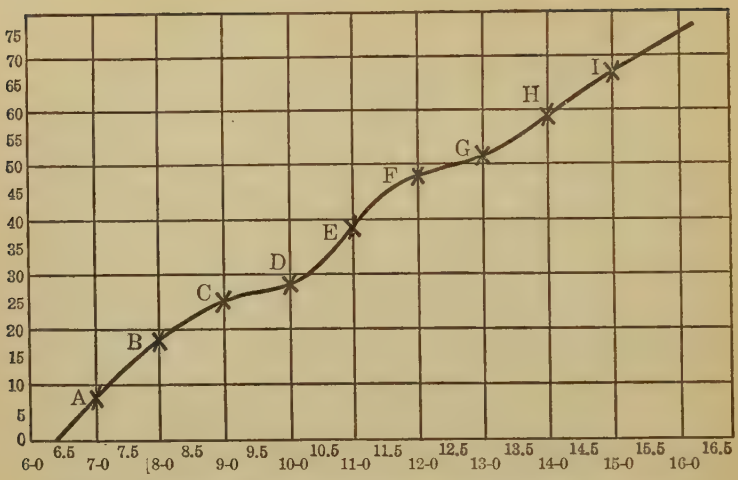


FIG. 75. Drawing the norm line on the transmutation sheet

When the table has been prepared with the scores in a column, it is necessary to consult the transmutation sheet and transfer the age norms for each score. The line for a score of 1 crosses the norm line on the age line 6–6. This age should then be placed in the table opposite the score of 1. The line for the score of 2 crosses the norm line on the age line 6–7, which should also be transferred.

Fig. 76 is an enlargement of the section of the transmutation sheet containing the norm line from the raw score of 0 to 18 and from the ages 6–6 to 8–0, and shows how each of the intervening scores and ages is determined by finding the points where the raw-score and age lines intersect with the norm line.

When all the intersections of the raw-score and age lines with the norm line have been determined and the points transferred to the age table, it will have the appearance of Table XXI on page 218.

Thereafter, to change a raw score in this test into an age

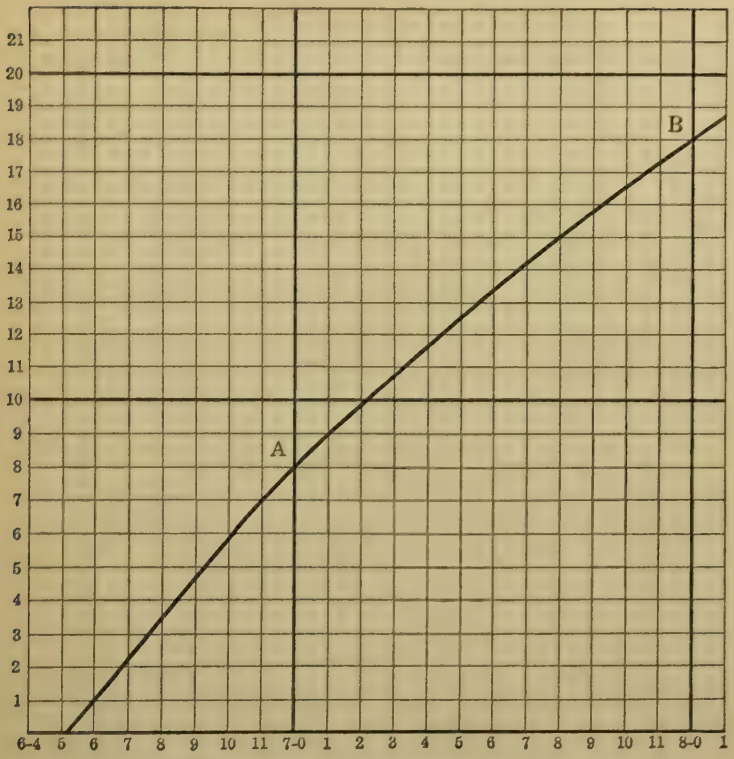


FIG. 76. Enlargement of portion of the norm line

score it is necessary only to locate the raw score on the table and find its appropriate age score.

Making an age-grade table. In using the age and grade norms it is frequently necessary to transpose raw scores into either age or grade scores, or to transpose grade scores into age scores, and the like. To do this, it is necessary to make an age-grade table for all the possible raw scores of the test. There are three ways in which such a table can be made up, depending on the data that may be available: when the norms

TABLE XXI. COMPLETED RAW-SCORE-AGE TABLE

RAW SCORE	AGE	RAW SCORE	AGE
1	6–6	38	10–11
2	6–7	39	11–0
3	6–8	40	11–1
4	6–9	41	11–2
5	6–9	42	11–3
6	6–10	43	11–5
7	6–11	44	11–6
8	7–0	45	11–8
9	7–1	46	11–9
10	7–2	47	11–11
11	7–4	48	12–0
12	7–5	49	12–3
13	7–6	50	12–6
14	7–7	51	12–9
15	7–8	52	13–0
16	7–9	53	13–2
17	7–11	54	13–4
18	8–0	55	13–5
19	8–2	56	13–7
20	8–3	57	13–9
21	8–5	58	13–10
22	8–6	59	14–0
23	8–8	60	14–2
24	8–10	61	14–3
25	9–0	62	14–5
26	9–3	63	14–6
27	9–7	64	14–8
28	10–0	65	14–9
29	10–1	66	14–11
30	10–3	67	15–0
31	10–4	68	15–2
32	10–5	69	15–4
33	10–6	70	15–6
34	10–8	71	15–8
35	10–9	72	15–9
36	10–10	73	15–10
37	10–10	74	16–0

of the test are given in both age norms and grade norms, when only age norms are given, and when only grade norms are available. The kind of norms available should be found and then the proper procedure should be selected.

Making an age-grade table when both grade and age norms are given. When both grade and age norms are given the following procedure is possible.

STEP 1. *Making a raw-score-grade table*

The construction of a grade table from a set of raw scores and a set of grade norms has already been described on pages 203–211. This consists in making a transmutation sheet, placing the grade norms and raw scores on the sheet, and reading directly from the sheet the various grade scores for the various raw scores.

In the illustration to be used the grade norms are given as in Table XXII, which is a list of grade norms for a test determined at the end of March.

TABLE XXII. GRADE NORMS ON A TEST TAKEN FROM THE DIRECTION MANUAL FOR THE TEST (NORMS AS OF END OF MARCH)

NORM FOR	RAW SCORE
Grade I	38
Grade II	68
Grade III	90
Grade IV	104
Grade V	115
Grade VI	125
Grade VII	133
Grade VIII	141

In making out a raw-score-grade table and placing the points on the transmutation sheet, it should be remembered that the grade norm as given on the table for Grade I is the grade norm for that grade at the end of March. This means that since March is the seventh month in the school year, the raw score of 38 is the norm for the grade of 1.7, the raw score of 68 is the norm for grade 2.7, and so on.

When the transmutation table has been completed for these norms, it has the appearance of Table XXIII on page 220.

TABLE XXIII. COMPLETED RAW-SCORE-GRADE TABLE

RAW SCORE	GRADE SCORE	RAW SCORE	GRADE SCORE
6	1.1	109	5.1
12	1.2	110	5.2
18	1.3	111	5.3
22	1.4	112	5.4
27	1.5	113	5.5
33	1.6	114	5.6
38	1.7	115	5.7
41	1.8	116	5.8
45	1.9	117	5.9
49	2.0	118	6.0
52	2.1	119	6.1
56	2.2	120	6.2
59	2.3	121	6.3
62	2.4	122	6.4
64	2.5	123	6.5
67	2.6	124	6.6
68	2.7	125	6.7
71	2.8	126	6.8
74	2.9	127	6.9
76	3.0	128	7.0
78	3.1	129	7.1
81	3.2	129	7.2
83	3.3	130	7.3
85	3.4	131	7.4
87	3.5	132	7.5
88	3.6	132	7.6
90	3.7	133	7.7
91	3.8	134	7.8
93	3.9	135	7.9
94	4.0	136	8.0
96	4.1	137	8.1
97	4.2	138	8.2
99	4.3	138	8.3
100	4.4	139	8.4
102	4.5	140	8.5
103	4.6	140	8.6
104	4.7	141	8.7
105	4.8	142	8.8
106	4.9	143	8.9
108	5.0	144	9.0

STEP 2. *Making a raw-score-age table*

The construction of an age table from the raw scores on a test has been described on pages 211–218. It is necessary to make a transmutation sheet, placing upon it the age norms taken from the direction manual of a test and, having drawn the norm line, reading directly from the sheet the age norms for the various raw scores.

In the test manual which is used for this illustration the age norms are given as in Table XXIV, which shows the age norms for the test.

TABLE XXIV. AGE NORMS FOR THE SAME TEST AS TABLE XXII AS TAKEN FROM THE MANUAL OF DIRECTIONS FOR THE TEST

AGE	RAW SCORE
7–0	38
8–0	67
9–0	87
10–0	100
11–0	111
12–0	121
13–0	129
14–0	137

After completing the transmutation table and recording the various age scores for the raw scores, the age table has the appearance of Table XXV. Here it will be noted that the scores have been placed only with the even ages. Thus scores of 7, 8, 9, 10, and 11 would be counted as age 6–5, scores 13, 14, 15, 16, and 17 as age 6–6, and so on.

STEP 3. *Combining the raw-score-grade and raw-score-age tables into one age-grade table*

In the two tables which have been found through Steps 1 and 2 there is a common element, the raw scores. In one table a grade score has been found for the raw scores; in the other an age score has been found. Matching up the age and grade scores for equivalent raw scores will give a completed grade-age table. This has been done, with the result shown in Table XXVI, where the age scores and grade scores have been placed together for equivalent raw scores.

TABLE XXV. COMPLETED RAW-SCORE-AGE TABLE

SCORE	AGE	SCORE	AGE	SCORE	AGE
1		49	7–4	97	9–9
2		50		98	
3		51	7–5	99	9–10
4		52	7–6	100	10–0
5		53		101	
6	6–5	54		102	10–1
7		55		103	10–2
8		56	7–7	104	10–3
9		57		105	10–5
10		58		106	10–6
11		59	7–8	107	10–7
12	6–6	60		108	10–8
13		61	7–9	109	10–9
14		62	7–10	110	10–10
15		63		111	11–0
16		64	7–11	112	11–1
17		65		113	11–2
18	6–7	66		114	11–3
19		67	8–0	115	11–4
20	6–8	68	8–1	116	11–6
21		69		117	11–7
22	6–9	70	8–2	118	11–8
23		71	8–3	119	11–10
24		72		120	11–11
25		73		121	12–0
26		74	8–4	122	12–2
27	6–10	75		123	12–3
28		76	8–5	124	12–4
29		77	8–6	125	12–5
30		78	8–7	126	12–7
31		79		127	12–8
32		80		128	12–10
33	6–11	81	8–8	129	13–0
34		82		130	13–1
35		83	8–9	131	13–3
36		84	8–10	132	13–4
37		85	8–11	133	13–6
38	7–0	86		134	13–8
39		87	9–0	135	13–9
40	7–1	88	9–1	136	13–10
41	7–2	89	9–2	137	14–0
42		90	9–3	138	14–1
43		91	9–4	139	14–3
44		92		140	14–5
45	7–3	93	9–5	141	14–7
46		94	9–6	142	14–9
47		95	9–7	143	14–10
48		96	9–8	144	14–11

TABLE XXVI. COMPLETED AGE-GRADE TABLE DERIVED FROM BOTH AGE AND GRADE NORMS

RAW SCORE	GRADE SCORE	AGE SCORE	RAW SCORE	GRADE SCORE	AGE SCORE
6	1.1	6–5	109	5.1	10–9
12	1.2	6–6	110	5.2	10–10
18	1.3	6–7	111	5.3	11–0
22	1.4	6–9	112	5.4	11–1
27	1.5	6–10	113	5.5	11–2
33	1.6	6–11	114	5.6	11–3
38	1.7	7–0	115	5.7	11–4
41	1.8	7–2	116	5.8	11–6
45	1.9	7–3	117	5.9	11–7
49	2.0	7–4	118	6.0	11–8
52	2.1	7–6	119	6.1	11–10
56	2.2	7–7	120	6.2	11–11
59	2.3	7–8	121	6.3	12–0
62	2.4	7–10	122	6.4	12–2
64	2.5	7–11	123	6.5	12–3
67	2.6	8–0	124	6.6	12–4
68	2.7	8–1	125	6.7	12–5
71	2.8	8–3	126	6.8	12–7
74	2.9	8–4	127	6.9	12–8
76	3.0	8–5	128	7.0	12–10
78	3.1	8–7	129	7.2	13–0
81	3.2	8–8	130	7.3	13–1
83	3.3	8–9	131	7.4	13–3
85	3.4	8–11	132	7.5	13–4
87	3.5	9–0	133	7.7	13–6
88	3.6	9–1	134	7.8	13–8
90	3.7	9–3	135	7.9	13–9
91	3.8	9–4	136	8.0	13–10
93	3.9	9–5	137	8.1	14–0
94	4.0	9–6	138	8.3	14–1
96	4.1	9–8	139	8.4	14–3
97	4.2	9–9	140	8.6	14–5
99	4.3	9–10	141	8.7	14–7
100	4.4	10–0	142	8.8	14–9
102	4.5	10–1	143	8.9	14–10
103	4.6	10–2	144	9.0	14–11
104	4.7	10–3			
105	4.8	10–5			
106	4.9	10–6			
108	5.0	10–9			

Making an age grade-table when only age norms are given. When only age norms are given for a test, a different means of constructing an age-grade table for the raw scores becomes necessary.

STEP 1. *Making a raw-score-age table*

The first step is to construct a raw-score-age table as has been described earlier in this chapter. If it be assumed that for the test which has been used here for illustration only age norms were given, the age table previously described can be used for illustration. This raw-score-age table is given in Table XXV.

STEP 2. *Determining the average age of children entering the first grade*

The average age of school children upon entrance to first grade varies in school systems according to the regulations governing entrance to school. Age norms as they are given in most tests are composite norms for children coming from different communities under different systems of restrictions regarding entrance to school. As will be shown, this sometimes reacts somewhat unfairly on the interpretations which are given to the achievement of pupils where the school system does not approximate the average age of other communities in its regulations for school entrance. For this reason an age-grade table constructed from age norms only is frequently better suited to a school system than the age-grade table made up of both age and grade norms taken for the test.

In some school systems children must be six years old before they can enter the first grade in September, in other systems they must be six years old before the first of October, and in still others they must be within six months of six years old. This means that the average age of children entering school varies from six years and six months in the systems where entrance is on the basis of having passed the sixth

birthday to six years and no months in the systems where they are allowed to enter when they are within six months of six years old.

In the illustration to be used here it is assumed that the pupils may enter first grade when they are within two months of six years old or if they are to be six years old before the first of November. The average age of such pupils is to be found by consulting Table XXVII, which gives the average ages of pupils at the beginning of first grade according to the possibilities for their entrance.

TABLE XXVII. SHOWING AVERAGE AGE AT ENTRANCE TO SCHOOL OF PUPILS UNDER DIFFERENT ENTRANCE CONDITIONS

When pupils may enter school within six months of six years old, the average age is 6–0
When pupils may enter school within five months of six years old, the average age is 6–1
When pupils may enter school within four months of six years old, the average age is 6–2
When pupils may enter school within three months of six years old, the average age is 6–3
When pupils may enter school within two months of six years old, the average age is 6–4
When pupils may enter school within one month of six years old, the average age is 6–5
When pupils may enter school at six years old, the average age is 6–6
When pupils may enter school at six years and one month old, the average age is 6–7
When pupils may enter school at six years and two months old, the average age is 6–8

STEP 3. *Transferring grade scores to the raw-score-age table in terms of the table here given*

Tables XXVIIIA and XXVIIIB are age-grade transmutation tables. Table XXVIIIA is given in years and months; Table XXVIIIB expresses the same material in months only. They include nine age tables ranging from an average age at entrance to first grade of six years and no months to six years

and eight months. It will be noticed that there are thirteen months in age equivalent to the ten months of progress in the school year. This is because it has been found that the average child progresses through school at a rate slightly less than one grade a year. This can be stated otherwise as meaning that in general there is greater retardation than acceleration in school.[1]

The age-grade transmutation table can be used as follows: select the proper age table, which is given in one of the columns labeled at the top from 6–0 to 6–8, according to the average age of pupils upon entrance to first grade, and then transfer to the raw-score-age table.

In the illustration here used the raw-score-age table is as given in Table XXV. In the system where the tests were given, let it be assumed that the pupils may enter school if they are within two months of six years old. Table XXVII gives the average age of pupils at entrance to first grade in this school system as 6–4. The age column then to be used in Table XXVIII A is the fifth column, which is labeled "6–4."

In Table XXV a score of 6 has an age score of 6–5; in Table XXVIII A, fifth column, an age score of 6–5 is a grade level of 1.1. In Table XXV a score of 12 has an age score of 6–6; in Table XXVIII A, fifth column, an age score of 6–6 has a grade score of 1.1. In Table XXV a score of 18 has an age score of 6–7; in Table XXVIII A, fifth column, an age score of 6–7 has a grade score of 1.2.

The process here started should be continued, transferring from Table XXVIII A the appropriate grade score found for each age score on the raw-score-age table until all are transferred. The completed table will resemble Table XXIX, where the derived grade scores are shown in the last column.

[1] See W. A. McCall's *How to Measure in Education* (The Macmillan Company, 1922), pp. 33–34.

Table XXVIII A. Age-Grade Transmutation Tables

Grade Score	Average-Age Scores Dependent on Age at Entering First Grade								
	6–0	6–1	6–2	6–3	6–4	6–5	6–6	6–7	6–8
1.0	6–0	6–1	6–2	6–3	6–4	6–5	6–6	6–7	6–8
1.1	*6–1*	*6–2*	*6–3*	*6–4*	*6–5*	*6–6*	*6–7*	*6–8*	*6–9*
	6–2	6–3	6–4	6–5	6–6	6–7	6–8	6–9	6–10
1.2	6–3	6–4	6–5	6–6	6–7	6–8	6–9	6–10	6–11
1.3	6–4	6–5	6–6	6–7	6–8	6–9	6–10	6–11	7–0
1.4	*6–5*	*6–6*	*6–7*	*6–8*	*6–9*	*6–10*	*6–11*	*7–0*	*7–1*
	6–6	6–7	6–8	6–9	6–10	6–11	7–0	7–1	7–2
1.5	6–7	6–8	6–9	6–10	6–11	7–0	7–1	7–2	7–3
1.6	6–8	6–9	6–10	6–11	7–0	7–1	7–2	7–3	7–4
1.7	6–9	6–10	6–11	7–0	7–1	7–2	7–3	7–4	7–5
	6–10	*6–11*	*7–0*	*7–1*	*7–2*	*7–3*	*7–4*	*7–5*	*7–6*
1.8	6–11	7–0	7–1	7–2	7–3	7–4	7–5	7–6	7–7
1.9	7–0	7–1	7–2	7–3	7–4	7–5	7–6	7–7	7–8
2.0	7–1	7–2	7–3	7–4	7–5	7–6	7–7	7–8	7–9
2.1	*7–2*	*7–3*	*7–4*	*7–5*	*7–6*	*7–7*	*7–8*	*7–9*	*7–10*
	7–3	7–4	7–5	7–6	7–7	7–8	7–9	7–10	7–11
2.2	7–4	7–5	7–6	7–7	7–8	7–9	7–10	7–11	8–0
2.3	7–5	7–6	7–7	7–8	7–9	7–10	7–11	8–0	8–1
2.4	*7–6*	*7–7*	*7–8*	*7–9*	*7–10*	*7–11*	*8–0*	*8–1*	*8–2*
	7–7	7–8	7–9	7–10	7–11	8–0	8–1	8–2	8–3
2.5	7–8	7–9	7–10	7–11	8–0	8–1	8–2	8–3	8–4
2.6	7–9	7–10	7–11	8–0	8–1	8–2	8–3	8–4	8–5
2.7	7–10	7–11	8–0	8–1	8–2	8–3	8–4	8–5	8–6
	7–11	*8–0*	*8–1*	*8–2*	*8–3*	*8–4*	*8–5*	*8–6*	*8–7*
2.8	8–0	8–1	8–2	8–3	8–4	8–5	8–6	8–7	8–8
2.9	8–1	8–2	8–3	8–4	8–5	8–6	8–7	8–8	8–9
3.0	8–2	8–3	8–4	8–5	8–6	8–7	8–8	8–9	8–10
3.1	*8–3*	*8–4*	*8–5*	*8–6*	*8–7*	*8–8*	*8–9*	*8–10*	*8–11*
	8–4	8–5	8–6	8–7	8–8	8–9	8–10	8–11	9–0
3.2	8–5	8–6	8–7	8–8	8–9	8–10	8–11	9–0	9–1
3.3	8–6	8–7	8–8	8–9	8–10	8–11	9–0	9–1	9–2
3.4	*8–7*	*8–8*	*8–9*	*8–10*	*8–11*	*9–0*	*9–1*	*9–2*	*9–3*
	8–8	8–9	8–10	8–11	9–0	9–1	9–2	9–3	9–4
3.5	8–9	8–10	8–11	9–0	9–1	9–2	9–3	9–4	9–5

TABLE. XXVIII A. AGE-GRADE TRANSMUTATION TABLES (CONTINUED)

GRADE SCORE	AVERAGE-AGE SCORES DEPENDENT ON AGE AT ENTERING FIRST GRADE								
	6-0	6-1	6-2	6-3	6-4	6-5	6-6	6-7	6-8
3.6	8–10	8–11	9–0	9–1	9–2	9–3	9–4	9–5	9–6
3.7	8–11	9–0	9–1	9–2	9–3	9–4	9–5	9–6	9–7
	9–0	*9–1*	*9–2*	*9–3*	*9–4*	*9–5*	*9–6*	*9–7*	*9–8*
3.8	9–1	9–2	9–3	9–4	9–5	9–6	9–7	9–8	9–9
3.9	9–2	9–3	9–4	9–5	9–6	9–7	9–8	9–9	9–10
4.0	9–3	9–4	9–5	9–6	9–7	9–8	9–9	9–10	9–11
4.1	*9–4*	*9–5*	*9–6*	*9–7*	*9–8*	*9–9*	*9–10*	*9–11*	*10–0*
	9–5	9–6	9–7	9–8	9–9	9–10	9–11	10–0	10–1
4.2	9–6	9–7	9–8	9–9	9–10	9–11	10–0	10–1	10–2
4.3	9–7	9–8	9–9	9–10	9–11	10–0	10–1	10–2	10–3
4.4	*9–8*	*9–9*	*9–10*	*9–11*	*10–0*	*10–1*	*10–2*	*10–3*	*10–4*
	9–9	9–10	9–11	10–0	10–1	10–2	10–3	10–4	10–5
4.5	9–10	9–11	10–0	10–1	10–2	10–3	10–4	10–5	10–6
4.6	9–11	10–0	10–1	10–2	10–3	10–4	10–5	10–6	10–7
4.7	10–0	10–1	10–2	10–3	10–4	10–5	10–6	10–7	10–8
	10–1	*10–2*	*10–3*	*10–4*	*10–5*	*10–6*	*10–7*	*10–8*	*10–9*
4.8	10–2	10–3	10–4	10–5	10–6	10–7	10–8	10–9	10–10
4.9	10–3	10–4	10–5	10–6	10–7	10–8	10–9	10–10	10–11
5.0	10–4	10–5	10–6	10–7	10–8	10–9	10–10	10–11	11–0
5.1	*10–5*	*10–6*	*10–7*	*10–8*	*10–9*	*10–10*	*10–11*	*11–0*	*11–1*
	10–6	10–7	10–8	10–9	10–10	10–11	11–0	11–1	11–2
5.2	10–7	10–8	10–9	10–10	10–11	11–0	11–1	11–2	11–3
5.3	10–8	10–9	10–10	10–11	11–0	11–1	11–2	11–3	11–4
5.4	*10–9*	*10–10*	*10–11*	*11–0*	*11–1*	*11–2*	*11–3*	*11–4*	*11–5*
	10–10	10–11	11–0	11–1	11–2	11–3	11–4	11–5	11–6
5.5	10–11	11–0	11–1	11–2	11–3	11–4	11–5	11–6	11–7
5.6	11–0	11–1	11–2	11–3	11–4	11–5	11–6	11–7	11–8
5.7	11–1	11–2	11–3	11–4	11–5	11–6	11–7	11–8	11–9
	11–2	*11–3*	*11–4*	*11–5*	*11–6*	*11–7*	*11–8*	*11–9*	*11–10*
5.8	11–3	11–4	11–5	11–6	11–7	11–8	11–9	11–10	11–11
5.9	11–4	11–5	11–6	11–7	11–8	11–9	11–10	11–11	12–0
6.0	11–5	11–6	11–7	11–8	11–9	11–10	11–11	12–0	12–1
6.1	*11–6*	*11–7*	*11–8*	*11–9*	*11–10*	*11–11*	*12–0*	*12–1*	*12–2*
	11–7	11–8	11–9	11–10	11–11	12–0	12–1	12–2	12–3

TABLE XXVIII A. AGE-GRADE TRANSMUTATION TABLES (CONTINUED)

GRADE SCORE	AVERAGE-AGE SCORES DEPENDENT ON AGE AT ENTERING FIRST GRADE								
	6–0	6–1	6–2	6–3	6–4	6–5	6–6	6–7	6–8
6.2	11–8	11–9	11–10	11–11	12–0	12–1	12–2	12–3	12–4
6.3	11–9	11–10	11–11	12–0	12–1	12–2	12–3	12–4	12–5
6.4	*11–10*	*11–11*	*12–0*	*12–1*	*12–2*	*12–3*	*12–4*	*12–5*	*12–6*
	11–11	12–0	12–1	12–2	12–3	12–4	12–5	12–6	12–7
6.5	12–0	12–1	12–2	12–3	12–4	12–5	12–6	12–7	12–8
6.6	12–1	12–2	12–3	12–4	12–5	12–6	12–7	12–8	12–9
6.7	12–2	12–3	12–4	12–5	12–6	12–7	12–8	12–9	12–10
	12–3	*12–4*	*12–5*	*12–6*	*12–7*	*12–8*	*12–9*	*12–10*	*12–11*
6.8	12–4	12–5	12–6	12–7	12–8	12–9	12–10	12–11	13–0
6.9	12–5	12–6	12–7	12–8	12–9	12–10	12–11	13–0	13–1
7.0	12–6	12–7	12–8	12–9	12–10	12–11	13–0	13–1	13–2
7.1	*12–7*	*12–8*	*12–9*	*12–10*	*12–11*	*13–0*	*13–1*	*13–2*	*13–3*
	12–8	12–9	12–10	12–11	13–0	13–1	13–2	13–3	13–4
7.2	12–9	12–10	12–11	13–0	13–1	13–2	13–3	13–4	13–5
7.3	12–10	12–11	13–0	13–1	13–2	13–3	13–4	13–5	13–6
7.4	*12–11*	*13–0*	*13–1*	*13–2*	*13–3*	*13–4*	*13–5*	*13–6*	*13–7*
	13–0	13–1	13–2	13–3	13–4	13–5	13–6	13–7	13–8
7.5	13–1	13–2	13–3	13–4	13–5	13–6	13–7	13–8	13–9
7.6	13–2	13–3	13–4	13–5	13–6	13–7	13–8	13–9	13–10
7.7	13–3	13–4	13–5	13–6	13–7	13–8	13–9	13–10	13–11
	13–4	*13–5*	*13–6*	*13–7*	*13–8*	*13–9*	*13–10*	*13–11*	*14–0*
7.8	13–5	13–6	13–7	13–8	13–9	13–10	13–11	14–0	14–1
7.9	13–6	13–7	13–8	13–9	13–10	13–11	14–0	14–1	14–2
8.0	13–7	13–8	13–9	13–10	13–11	14–0	14–1	14–2	14–3
8.1	*13–8*	*13–9*	*13–10*	*13–11*	*14–0*	*14–1*	*14–2*	*14–3*	*14–4*
	13–9	13–10	13–11	14–0	14–1	14–2	14–3	14–4	14–5
8.2	13–10	13–11	14–0	14–1	14–2	14–3	14–4	14–5	14–6
8.3	13–11	14–0	14–1	14–2	14–3	14–4	14–5	14–6	14–7
8.4	*14–0*	*14–1*	*14–2*	*14–3*	*14–4*	*14–5*	*14–6*	*14–7*	*14–8*
	14–1	14–2	14–3	14–4	14–5	14–6	14–7	14–8	14–9
8.5	14–2	14–3	14–4	14–5	14–6	14–7	14–8	14–9	14–10
8.6	14–3	14–4	14–5	14–6	14–7	14–8	14–9	14–10	14–11
8.7	14–4	14–5	14–6	14–7	14–8	14–9	14–10	14–11	15–0
	14–5	*14–6*	*14–7*	*14–8*	*14–9*	*14–10*	*14–11*	*15–0*	*15–1*
8.8	14–6	14–7	14–8	14–9	14–10	14–11	15–0	15–1	15–2
8.9	14–7	14–8	14–9	14–10	14–11	15–0	15–1	15–2	15–3

TABLE XXVIII A. AGE-GRADE TRANSMUTATION TABLES (CONCLUDED)

GRADE SCORE	AVERAGE-AGE SCORES DEPENDENT ON AGE AT ENTERING FIRST GRADE								
	6–0	6–1	6–2	6–3	6–4	6–5	6–6	6–7	6–8
9.0	14–8	14–9	14–10	14–11	15–0	15–1	15–2	15–3	15–4
9.1	*14–9*	*14–10*	*14–11*	*15–0*	*15–1*	*15–2*	*15–3*	*15–4*	*15–5*
	14–10	14–11	15–0	15–1	15–2	15–3	15–4	15–5	15–6
9.2	14–11	15–0	15–1	15–2	15–3	15–4	15–5	15–6	15–7
9.3	15–0	15–1	15–2	15–3	15–4	15–5	15–6	15–7	15–8
9.4	*15–1*	*15–2*	*15–3*	*15–4*	*15–5*	*15–6*	*15–7*	*15–8*	*15–9*
	15–2	15–3	15–4	15–5	15–6	15–7	15–8	15–9	15–10
9.5	15–3	15–4	15–5	15–6	15–7	15–8	15–9	15–10	15–11
9.6	15–4	15–5	15–6	15–7	15–8	15–9	15–10	15–11	16–0
9.7	15–5	15–6	15–7	15–8	15–9	15–10	15–11	16–0	16–1
	15–6	*15–7*	*15–8*	*15–9*	*15–10*	*15–11*	*16–0*	*16–1*	*16–2*
9.8	15–7	15–8	15–9	15–10	15–11	16–0	16–1	16–2	16–3
9.9	15–8	15–9	15–10	15–11	16–0	16–1	16–2	16–3	16–4

TABLE XXVIII B. AGE-GRADE TRANSMUTATION TABLES

GRADE SCORE	AVERAGE-AGE SCORES DEPENDENT ON AGE AT ENTERING FIRST GRADE								
	6–0	6–1	6–2	6–3	6–4	6–5	6–6	6–7	6–8
1.0	72	73	74	75	76	77	78	79	80
1.1	*73*	*74*	*75*	*76*	*77*	*78*	*79*	*80*	*81*
	74	75	76	77	78	79	80	81	82
1.2	75	76	77	78	79	80	81	82	83
1.3	76	77	78	79	80	81	82	83	84
1.4	*77*	*78*	*79*	*80*	*81*	*82*	*83*	*84*	*85*
	78	79	80	81	82	83	84	85	86
1.5	79	80	81	82	83	84	85	86	87
1.6	80	81	82	83	84	85	86	87	88
1.7	81	82	83	84	85	86	87	88	89
	82	*83*	*84*	*85*	*86*	*87*	*88*	*89*	*90*
1.8	83	84	85	86	87	88	89	90	91
1.9	84	85	86	87	88	89	90	91	92
2.0	85	86	87	88	89	90	91	92	93

TABLE XXVIII B. AGE-GRADE TRANSMUTATION TABLES (CONTINUED)

GRADE SCORE	AVERAGE-AGE SCORES DEPENDENT ON AGE AT ENTERING FIRST GRADE								
	6–0	6–1	6–2	6–3	6–4	6–5	6–6	6–7	6–8
2.1	*86*	*87*	*88*	*89*	*90*	*91*	*92*	*93*	*94*
	87	88	89	90	91	92	93	94	95
2.2	88	89	90	91	92	93	94	95	96
2.3	89	90	91	92	93	94	95	96	97
2.4	*90*	*91*	*92*	*93*	*94*	*95*	*96*	*97*	*98*
	91	92	93	94	95	96	97	98	99
2.5	92	93	94	95	96	97	98	99	100
2.6	93	94	95	96	97	98	99	100	101
2.7	94	95	96	97	98	99	100	101	102
	95	*96*	*97*	*98*	*99*	*100*	*101*	*102*	*103*
2.8	96	97	98	99	100	101	102	103	104
2.9	97	98	99	100	101	102	103	104	105
3.0	98	99	100	101	102	103	104	105	106
3.1	*99*	*100*	*101*	*102*	*103*	*104*	*105*	*106*	*107*
	100	101	102	103	104	105	106	107	108
3.2	101	102	103	104	105	106	107	108	109
3.3	102	103	104	105	106	107	108	109	110
3.4	*103*	*104*	*105*	*106*	*107*	*108*	*109*	*110*	*111*
	104	105	106	107	108	109	110	111	112
3.5	105	106	107	108	109	110	111	112	113
3.6	106	107	108	109	110	111	112	113	114
3.7	107	108	109	110	111	112	113	114	115
	108	*109*	*110*	*111*	*112*	*113*	*114*	*115*	*116*
3.8	109	110	111	112	113	114	115	116	117
3.9	110	111	112	113	114	115	116	117	118
4.0	111	112	113	114	115	116	117	118	119
4.1	*112*	*113*	*114*	*115*	*116*	*117*	*118*	*119*	*120*
	113	114	115	116	117	118	119	120	121
4.2	114	115	116	117	118	119	120	121	122
4.3	115	116	117	118	119	120	121	122	123
4.4	*116*	*117*	*118*	*119*	*120*	*121*	*122*	*123*	*124*
	117	118	119	120	121	122	123	124	125
4.5	118	119	120	121	122	123	124	125	126
4.6	119	120	121	122	123	124	125	126	127

TABLE XXVIII B. AGE-GRADE TRANSMUTATION TABLES (CONTINUED)

GRADE SCORE	AVERAGE-AGE SCORES DEPENDENT ON AGE AT ENTERING FIRST GRADE								
	6–0	6–1	6–2	6–3	6–4	6–5	6–6	6–7	6–8
4.7	120	121	122	123	124	125	126	127	128
	121	*122*	*123*	*124*	*125*	*126*	*127*	*128*	*129*
4.8	122	123	124	125	126	127	128	129	130
4.9	123	124	125	126	127	128	129	130	131
5.0	124	125	126	127	128	129	130	131	132
5.1	*125*	*126*	*127*	*128*	*129*	*130*	*131*	*132*	*133*
	126	127	128	129	130	131	132	133	134
5.2	127	128	129	130	131	132	133	134	135
5.3	128	129	130	131	132	133	134	135	136
5.4	*129*	*130*	*131*	*132*	*133*	*134*	*135*	*136*	*137*
	130	131	132	133	134	135	136	137	138
5.5	131	132	133	134	135	136	137	138	139
5.6	132	133	134	135	136	137	138	139	140
5.7	133	134	135	136	137	138	139	140	141
	134	*135*	*136*	*137*	*138*	*139*	*140*	*141*	*142*
5.8	135	136	137	138	139	140	141	142	143
5.9	136	137	138	139	140	141	142	143	144
6.0	137	138	139	140	141	142	143	144	145
6.1	*138*	*139*	*140*	*141*	*142*	*143*	*144*	*145*	*146*
	139	140	141	142	143	144	145	146	147
6.2	140	141	142	143	144	145	146	147	148
6.3	141	142	143	144	145	146	147	148	149
6.4	*142*	*143*	*144*	*145*	*146*	*147*	*148*	*149*	*150*
	143	144	145	146	147	148	149	150	151
6.5	144	145	146	147	148	149	150	151	152
6.6	145	146	147	148	149	150	151	152	153
6.7	146	147	148	149	150	151	152	153	154
	147	*148*	*149*	*150*	*151*	*152*	*153*	*154*	*155*
6.8	148	149	150	151	152	153	154	155	156
6.9	149	150	151	152	153	154	155	156	157
7.0	150	151	152	153	154	155	156	157	158
7.1	*151*	*152*	*153*	*154*	*155*	*156*	*157*	*158*	*159*
	152	153	154	155	156	157	158	159	160
7.2	153	154	155	156	157	158	159	160	161

TABLE XXVIII B. AGE-GRADE TRANSMUTATION TABLES (CONCLUDED)

GRADE SCORE	AVERAGE-AGE SCORES DEPENDENT ON AGE AT ENTERING FIRST GRADE								
	6–0	6–1	6–2	6–3	6–4	6–5	6–6	6–7	6–8
7.3	154	155	156	157	158	159	160	161	162
7.4	*155*	*156*	*157*	*158*	*159*	*160*	*161*	*162*	*163*
	156	157	158	159	160	161	162	163	164
7.5	157	158	159	160	161	162	163	164	165
7.6	158	159	160	161	162	163	164	165	166
7.7	159	160	161	162	163	164	165	166	167
	160	*161*	*162*	*163*	*164*	*165*	*166*	*167*	*168*
7.8	161	162	163	164	165	166	167	168	169
7.9	162	163	164	165	166	167	168	169	170
8.0	163	164	165	166	167	168	169	170	171
8.1	*164*	*165*	*166*	*167*	*168*	*169*	*170*	*171*	*172*
	165	166	167	168	169	170	171	172	173
8.2	166	167	168	169	170	171	172	173	174
8.3	167	168	169	170	171	172	173	174	175
8.4	*168*	*169*	*170*	*171*	*172*	*173*	*174*	*175*	*176*
	169	170	171	172	173	174	175	176	177
8.5	170	171	172	173	174	175	176	177	178
8.6	171	172	173	174	175	176	177	178	179
8.7	172	173	174	175	176	177	178	179	180
	173	*174*	*175*	*176*	*177*	*178*	*179*	*180*	*181*
8.8	174	175	176	177	178	179	180	181	182
8.9	175	176	177	178	179	180	181	182	183
9.0	176	177	178	179	180	181	182	183	184
9.1	*177*	*178*	*179*	*180*	*181*	*182*	*183*	*184*	*185*
	178	179	180	181	182	183	184	185	186
9.2	179	180	181	182	183	184	185	186	187
9.3	180	181	182	183	184	185	186	187	188
9.4	*181*	*182*	*183*	*184*	*185*	*186*	*187*	*188*	*189*
	182	183	184	185	186	187	188	189	190
9.5	183	184	185	186	187	188	189	190	191
9.6	184	185	186	187	188	189	190	191	192
9.7	185	186	187	188	189	190	191	192	193
	186	*187*	*188*	*189*	*190*	*191*	*192*	*193*	*194*
9.8	187	188	189	190	191	192	193	194	195
9.9	188	189	190	191	192	193	194	195	196

TABLE XXIX. COMPLETED AGE-GRADE TABLE CONSTRUCTED FROM AGE NORMS ONLY (BASED ON AVERAGE ENTERING AGE OF 6–4)

RAW SCORE	AGE SCORE	GRADE SCORE	RAW SCORE	AGE SCORE	GRADE SCORE
6	6–5	1.1	102	10–1	4.4
12	6–6	1.1	103	10–2	4.5
18	6–7	1.2	104	10–3	4.6
20	6–8	1.3	105	10–5	4.7
22	6–9	1.4	106	10–6	4.8
27	6–10	1.4	107	10–7	4.9
33	6–11	1.5	108	10–8	5.0
38	7–0	1.6	109	10–9	5.1
40	7–1	1.7	110	10–10	5.1
41	7–2	1.7	111	11–0	5.3
45	7–3	1.8	112	11–1	5.4
49	7–4	1.9	113	11–2	5.4
51	7–5	2.0	114	11–3	5.5
52	7–6	2.1	115	11–4	5.6
56	7–7	2.1	116	11–6	5.7
59	7–8	2.2	117	11–7	5.8
61	7–9	2.3	118	11–8	5.9
62	7–10	2.4	119	11–10	6.1
64	7–11	2.4	120	11–11	6.1
67	8–0	2.5	121	12–0	6.2
68	8–1	2.6	122	12–2	6.4
70	8–2	2.7	123	12–3	6.4
71	8–3	2.7	124	12–4	6.5
74	8–4	2.8	125	12–5	6.6
76	8–5	2.9	126	12–7	6.7
77	8–6	3.0	127	12–8	6.8
78	8–7	3.1	128	12–10	7.0
81	8–8	3.1	129	13–0	7.1
83	8–9	3.2	130	13–1	7.2
84	8–10	3.3	131	13–3	7.4
85	8–11	3.4	132	13–4	7.4
87	9–0	3.4	133	13–6	7.6
88	9–1	3.5	134	13–8	7.7
89	9–2	3.6	135	13–9	7.8
90	9–3	3.7	136	13–10	7.9
91	9–4	3.7	137	14–0	8.0
93	9–5	3.8	138	14–1	8.1
94	9–6	3.9	139	14–3	8.2
95	9–7	4.0	140	14–5	8.4
96	9–8	4.1	141	14–7	8.6
97	9–9	4.1	142	14–9	8.7
99	9–10	4.2	143	14–10	8.8
100	10–0	4.3	144	14–11	8.9

Making an age-grade table when only grade norms are given. In making an age-grade table when only grade norms are available the process described in the foregoing sections must be reversed, so far as the use of the age-grade transmutation table is concerned.

STEP 1. *Making a raw-score-grade table*

The process in making a raw-score-grade table has been described in detail on pages 203–211. The grade norms are taken from the direction manual of the test and are transmuted into grade scores for each raw score on the test. A tabulation of these grade scores for each raw score results in the raw-score-grade table, an example of which is shown in Table XXIII.

STEP 2. *Determining the average age of pupils entering the first grade*

The reasons for making this step and the process for so doing were described on page 224 of this chapter. By determining the age at which children are allowed to enter school and by consulting Table XXVII, the average can be found.

If it is assumed that there were no age norms given for the test used in Table XXIII, that the grade norms are determined as found there, and that the system in which the test was given allows pupils to enter school when they are within four months of being six years old, the age-grade table is easily made.

STEP 3. *Transferring appropriate age scores as given in Table XXVIII A*

The final step is to find in Table XXVIII A or Table XXVIII B, tables showing the transmutation of grade and age units in terms of nine different average ages of beginning school, the appropriate age scores for the grade scores given.

In the example used in Table XXIII, a raw score of 6 is equivalent to a grade score of 1.1. Table XXVIII A, third

column, shows that a grade score of 1.1 is an age score of 6–4.[1] Here it will be noted that there are two possible ages for the grade score of 1.1. The score given in plain type is preferable and should be used throughout the table when there is a choice; the italic figures are for use in finding grade scores from age scores, using the same table.

A raw score of 12 on Table XXIII is a grade score of 1.2. Table XXVIII A shows that a grade score of 1.2 is 6–5 in cases where the average age of entering first grade is 6–2.[2]

Continuing the same process for all the other grade scores in Table XXIII will result in a final completed age-grade table derived from grade norms only. This completed table is shown as Table XXX.

No confusion should result if it is noticed that in the three examples given for finding age-grade tables the three results appear to be different. This is because the three sets of data, although they are really the same, have been treated on different assumptions. In the first age-grade table it is assumed that both age and grade norms were available, whereas in the second it was assumed that the grade norms were derived from the supposition that children could enter school when within two months of six years old, and in the third that they could enter when within four months of six years old. The differences caused by these variations in the assumptions are revealed in the final tables.

Achievement measures in the T scale. To correct some of the difficulties that are present in the use of either the grade scale or the age scale, a third scale has been developed with a different set of units. This is called the T scale, so named by Dr. McCall, its originator. The achievement measure in the T scale is the T score. The T score is the score of a

[1] Found from Table XXVII.
[2] See Table XXVIII A, third column.

TABLE XXX. COMPLETED AGE-GRADE TABLE DERIVED WHEN GRADE NORMS ONLY WERE AVAILABLE

RAW SCORE	GRADE SCORE	AGE SCORE	RAW SCORE	GRADE SCORE	AGE SCORE
6	1.1	6–5	109	5.1	10–8
12	1.2	6–4	110	5.2	10–9
18	1.3	6–6	111	5.3	10–10
22	1.4	6–8	112	5.4	11–0
27	1.5	6–9	113	5.5	11–1
33	1.6	6–10	114	5.6	11–2
38	1.7	6–11	115	5.7	11–3
41	1.8	7–1	116	5.8	11–5
45	1.9	7–2	117	5.9	11–6
49	2.0	7–3	118	6.0	11–7
52	2.1	7–5	119	6.1	11–9
56	2.2	7–6	120	6.2	11–10
59	2.3	7–7	121	6.3	11–11
62	2.4	7–9	122	6.4	12–1
64	2.5	7–10	123	6.5	12–2
67	2.6	7–11	124	6.6	12–3
68	2.7	8–0	125	6.7	12–4
71	2.8	8–2	126	6.8	12–6
74	2.9	8–3	127	6.9	12–7
76	3.0	8–4	128	7.0	12–8
78	3.1	8–6	129	7.2	12–11
81	3.2	8–7	130	7.3	13–0
83	3.3	8–8	131	7.4	13–2
85	3.4	8–10	132	7.5	13–3
87	3.5	8–11	133	7.7	13–5
88	3.6	9–0	134	7.8	13–7
90	3.7	9–1	135	7.9	13–8
91	3.8	9–3	136	8.0	13–9
93	3.9	9–4	137	8.1	13–11
94	4.0	9–5	138	8.2	14–0
96	4.1	9–7	149	8.4	14–3
97	4.2	9–8	140	8.5	14–4
99	4.3	9–9	141	8.7	14–6
100	4.4	9–11	142	8.8	14–8
102	4.5	10–0	143	8.9	14–9
103	4.6	10–1	144	9.0	14–10
104	4.7	10–2			
105	4.8	10–4			
106	4.9	10–5			
108	5.0	10–6			

pupil on a T scale which is comparable to the placement of a pupil twelve years and six months old receiving the same score on the same scale.

Derivation of the T scale and meaning of a T score. In determining the T scale it was thought that some standard could be found which had more general meaning and more uniform significance than the standards of age and grade. This standard had necessarily to be in terms of the achievement of children and in terms of just as many children as possible, so that the averages of the standard would be as reliable as possible. To make the standard useful, it was decided to utilize for it the achievement of a group of pupils of a single age and to select that age as the one in which there were likely to be most pupils in the school grades. The age selected was twelve years old; that is, all children just over twelve years and just under thirteen years. This particular age is characteristic of school ages. Practically every child who is twelve years old is to be found somewhere in the elementary or junior-high-school grades, and these children are distributed from the first grade up.

Since there is such a wide distribution of twelve-year-old children throughout the school grades, the achievement of these pupils varies from work of the type done in the lower grades to work of the type done in the higher grades, and therefore the achievement of practically all the pupils in the elementary school can be compared with the work of some type of twelve-year-old child.

To represent the work of a twelve-year-old pupil in units which are standard for every test, the results of the work of twelve-year-old children were gathered and formed into a frequency surface. This frequency surface, similar in appearance to Fig. 63, p. 190, approximates closely a normal curve because of the many factors which contributed to its make-up.

When the frequency surface had been made, the next step in the derivation of the T unit was to determine the standard deviation from the mean of the curve. This deviation, commonly called the S.D., is represented as a distance on the base line of a curve and is determined by the degree to which the curve deviates from the mean of the curve. Since the normal curve theoretically never meets the base line, the S.D. distances on the base line can be produced indefinitely.[1] In constructing the T scale Dr. McCall utilized for his point of reference plus and minus 5 S.D. as the limits of his scale, with the mean zero S.D., as the mean of the achievement of twelve-year-old children.

Plus and minus 5 S.D. included 99.999971 per cent of twelve-year-old children (more than any teacher would ever need to utilize for a standard) and also represented quite practically a range of achievement which extended below and above the range of elementary-school work.

With the S.D. distances determined, the next step was to transpose the S.D. units found into units somewhat easier to handle, and this was done by transmuting the S.D. units directly into T units of exactly the same size as $\frac{1}{10}$ S.D. Minus 5 S.D., taken arbitrarily as the lowest point in the scale, was selected as the zero point of the new scale and was called 0. Each S.D. distance above -5 S.D. was divided into 10 units, so that the point on the scale equivalent to $+5$ S.D. became 100. The advantage in making this transmutation was that although the value of the S.D. scale was not changed a particle, the new scale eliminated decimal points as well as minus quantities. A further advantage was that the reference point, 50, was a quantity easy to remember and convenient to handle.

The derivation of the T scale from 0 to 100, as transmuted from the S.D. scale, is graphically shown in Fig. 77.

[1] See explanation on page 189.

It will be noted that the figures on the T scale exactly parallel those on the S.D. scale. They are, as a matter of fact, exactly the same thing under a new name and in a more easily readable form. Minus 5 S.D. becomes 0 T, — 4.9 S.D. becomes 1 T, and so on; whereas 0 S.D. becomes 50 T, 4.9 S.D. becomes 99 T, and 5.0 S.D. becomes 100 T.

Using the T table, it becomes possible to say that a child who has a T score of 50 has an achievement equal to that of an average twelve-year-old child in that field, that a child who receives a T score of 40 is 10 T below the mean of twelve-year-old children, and that a child who receives a T score of 60 is 10 T above the mean of twelve-year-old children. This difference can be expressed in various ways, as will be pointed out later.

Transmutation of raw scores into T scores. When T tables are not available for any particular test, it is necessary to transmute the raw scores for the test into T scores before the T-scale system can be used. There are certain tests, both achievement tests and intelligence tests, for which standard T tables have been constructed. These tables can be procured at nominal cost from the Teachers College Bureau of Publications, Columbia University, New York City. For other tests the following procedure for transmuting raw scores into T scores can be used.

Construction of a T Table for any test when grade norms are available. Some tests are published with grade norms only. Although a T table is probably more useful when made from age norms, one can be made from grade norms as follows.

STEP 1. *Constructing a raw-score-grade table*

The steps of the construction of a raw-score-grade table have been described on pages 203–211. They consist in listing the raw scores, making a transmutation sheet by placing the grade norms upon it, drawing the norm line, and

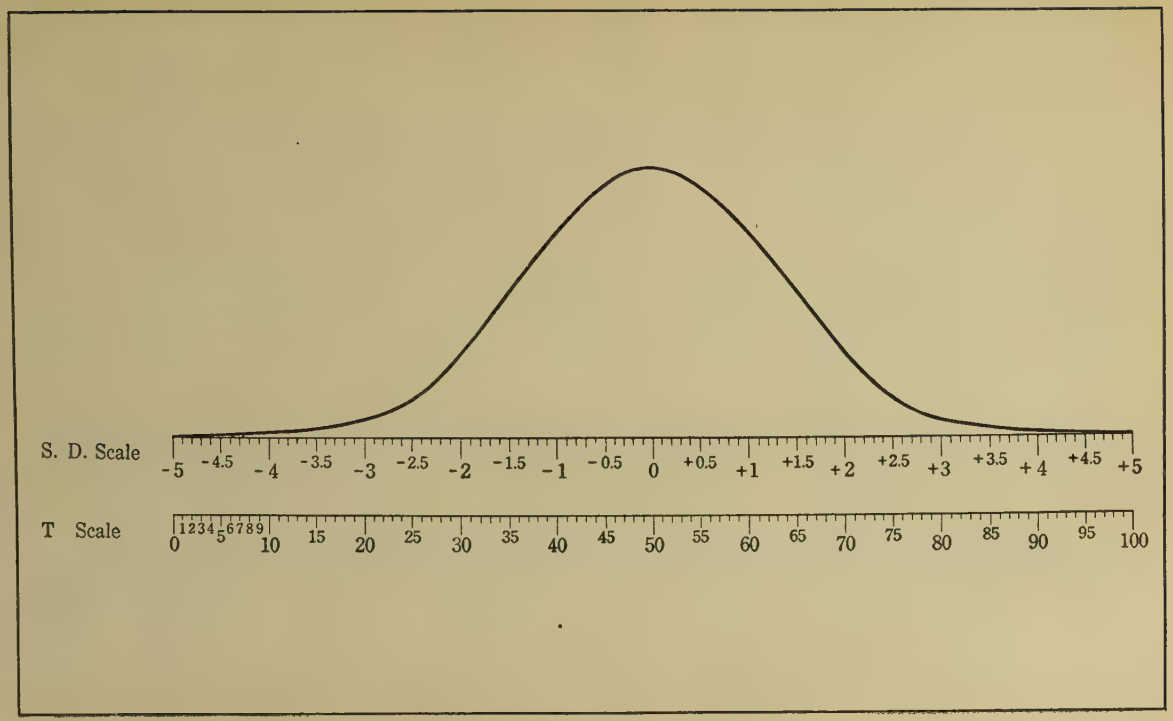


FIG. 77. Derivation of the T scale

transmuting the grade scores for the appropriate raw scores. When this process has been completed, the result will be as in the example in Table XXIII, p. 220.

STEP 2. *Transmuting grade scores into T scores according to Table XXXI*

The final step is the transferring of the appropriate T scores to the raw-score-grade table according to the data found in Table XXXI, which consists of a listing of age and grade scores, with the approximate T score for each. It was constructed by Dr. McCall and is here used by his permission.

Table XXXI is read by finding the grade score for any raw score and, reading horizontally across, finding the T score for the grade score.

According to Table XXIII a raw score of 6 has a grade score of 1.1. Table XXXI shows this grade score to be equivalent to a T score of 9. A grade score of 1.2 is equivalent to a T score of 10; a grade score of 1.3 has a T score of 11. The same process followed through for all the grade scores on Table XXIII will give a completed T table for that group of scores, as is shown in Table XXXII.

It should be noted that when a T table is once made up for any test, it is not necessary to remake it each time the test is given, unless there is some change in the norms. When the table is once made up, it is only necessary to read directly from it the T score for any raw score that may be met with. In the case of a T table such as the one given in Table XXXII any score for which a T score is not given should be given a T score equivalent to the next lower score which has a T rating; thus a raw score of 13, 14, 15, etc. would have T equivalents of the score 12, which is 10, and the scores 34, 35, 36, and 37 would have T equivalents of the score 33, which is 14.

TABLE XXXI.[1] TABLE FOR CONSTRUCTING APPROXIMATE T TABLES AND AGE TABLES AND FOR CONVERTING AGE SCORES INTO T SCORES AND T SCORES INTO AGE SCORES

T Score	Age Score	Grade Score	T Score	Age Score	Grade Score	T Score	Age Score	Grade Score
−30	3–0		13	6–11	1.5	56	14–2	8.0
−29	3–1		14	7–0	1.6	57	14–5	8.2
−28	3–2		15	7–1	1.7	58	14–8	8.4
−27	3–3		16	7–2	1.8	59	15–0	8.5
−26	3–5		17	7–3	1.9	60	15–5	8.6
−25	3–6		18	7–4	2.0	61	15–9	8.8
−24	3–7		19	7–5	2.1	62	16–1	9.0
−23	3–8		20	7–7	2.2	63	16–5	9.2
−22	3–10		21	7–8	2.3	64	16–10	9.4
−21	3–11		22	7–9	2.4	65	17–2	9.6
−20	4–0		23	7–10	2.5	66	17–6	9.8
−19	4–1		24	8–0	2.6	67	17–9	9.9
−18	4–2		25	8–1	2.7	68	18–0	10.0
−17	4–3		26	8–3	2.8	69	18–3	10.2
−16	4–5		27	8–4	2.9	70	18–6	10.4
−15	4–6		28	8–6	3.0	71	18–9	10.6
−14	4–7		29	8–7	3.1	72	19–0	10.8
−13	4–8		30	8–9	3.3	73	19–2	11.0
−12	4–10		31	8–10	3.4	74	19–5	11.2
−11	4–11		32	9–0	3.6	75	19–7	11.4
−10	5–0		33	9–2	3.8	76	19–10	11.6
− 9	5–1		34	9–4	4.0	77	20–0	11.7
− 8	5–2		35	9–5	4.2	78	20–3	11.8
− 7	5–3		36	9–7	4.4	79	20–5	11.9
− 6	5–4		37	9–8	4.6	80	20–8	12.0
− 5	5–5		38	10–0	4.8	81	20–10	12.2
− 4	5–7		39	10–2	4.9	82	21–1	12.4
− 3	5–8		40	10–4	5.0	83	21–3	12.6
− 2	5–10	0.0	41	10–6	5.2	84	21–6	12.8
− 1	5–11	0.1	42	10–8	5.4	85	21–8	13.0
0	6–0	0.2	43	10–11	5.6	86	21–11	13.2
1	6–1	0.3	44	11–2	5.8	87	22–1	13.4
2	6–2	0.4	45	11–4	6.0	88	22–4	13.6
3	6–2	0.5	46	11–6	6.2	89	22–6	13.7
4	6–3	0.6	47	11–9	6.4	90	22–8	13.8
5	6–3	0.7	48	12–0	6.6	91	22–10	14.0
6	6–4	0.8	49	12–2	6.7	92	23–0	14.2
7	6–5	0.9	50	12–6	6.8	93	23–2	14.4
8	6–6	1.0	51	12–9	7.0	94	23–4	14.6
9	6–7	1.1	52	13–0	7.2	95	23–6	14.8
10	6–8	1.2	53	13–3	7.4	96	23–8	15.0
11	6–9	1.3	54	13–7	7.6	97	23–10	15.2
12	6–10	1.4	55	13–10	7.8	98	24–0	15.4
						99	24–2	15.6
						100	24–4	15.8

[1] By permission of Teachers College Bureau of Publications.

TABLE XXXII. T TABLE CONSTRUCTED FROM RAW-SCORE-GRADE TABLE (TABLE XXIII)

RAW SCORE	GRADE SCORE	T SCORE	RAW SCORE	GRADE SCORE	T SCORE
6	1.1	9	109	5.1	40
12	1.2	10	110	5.2	41
18	1.3	11	111	5.3	41
22	1.4	12	112	5.4	42
27	1.5	13	113	5.5	42
33	1.6	14	114	5.6	43
38	1.7	15	115	5.7	43
41	1.8	16	116	5.8	44
45	1.9	17	117	5.9	44
49	2.0	18	118	6.0	45
52	2.1	19	119	6.1	45
56	2.2	20	120	6.2	46
59	2.4	21	121	6.3	46
62	2.4	22	122	6.4	47
64	2.5	23	123	6.5	47
67	2.6	24	124	6.6	48
68	2.7	25	125	6.7	49
71	2.8	26	126	6.8	50
74	2.9	27	127	6.9	50
76	3.0	28	128	7.0	51
78	3.1	29	129	7.2	52
81	3.2	29	130	7.3	52
83	3.3	30	131	7.4	53
85	3.4	31	132	7.5	53
87	3.5	31	133	7.7	54
88	3.6	32	134	7.8	55
90	3.7	32	135	7.9	55
91	3.8	33	136	8.0	56
93	3.9	33	137	8.1	56
94	4.0	34	138	8.2	57
96	4.1	34	139	8.4	58
97	4.2	35	140	8.5	59
99	4.3	35	141	8.7	60
100	4.4	36	142	8.8	61
102	4.5	36	143	8.9	61
103	4.6	37	144	9.0	62
104	4.7	37			
105	4.8	38			
106	4.9	39			
108	5.0	40			

Construction of a T table when age norms are given. When age norms are given, it is probable that the T table constructed from them will be somewhat more reliable than one constructed from grade norms, for the reason that there is probably more equivalence in the age levels of children in various communities than there is in their grade levels. The process is exactly the same as that described for grade norms.

STEP 1. *Constructing a raw-score-age table*

A raw-score-age table is constructed as described on pages 212–218, by listing the raw scores, making a transmutation sheet, placing on the sheet the points corresponding to the age norms, and, after constructing the age-norm line, transferring to the raw-score sheet the appropriate age scores indicated.

STEP 2. *Transferring to the raw-score-age table the T scores found from Table XXXI*

In Table XXXI will be found, in addition to the grade scores, a series of age scores corresponding to a series of T scores. The process in this step is to find the T scores which are appropriate for the age scores on the raw-score-age table of Step 1.

Table XXV, p. 222, is a completed raw-score-age table constructed as described above. A raw score of 6 is equivalent to an age score of 6–5. Table XXXI gives an age score of 6–5 as equivalent to a T score of 7. The next score on Table XXV is a score of 12, with an equivalent age score of 6–6; on Table XXXI this is equivalent to a T score of 8. The remaining age scores are transmuted in the same way and transferred to the raw-score-age table. Table XXXIII shows the completed T table resulting from this transmutation.

Chapter summary. There are many possible scales by which to measure the achievement of children. The three of greatest use to the teacher are the grade scale, the age scale, and the T scale.

TABLE XXXIII. T TABLE CONSTRUCTED FROM RAW-SCORE-AGE TABLE (TABLE XXV)

RAW SCORE	AGE SCORE	T SCORE	RAW SCORE	AGE SCORE	T SCORE
6	6–5	7	102	10–1	38
12	6–6	8	103	10–2	39
18	6–7	9	104	10–3	39
20	6–8	10	105	10–5	40
22	6–9	11	106	10–6	41
27	6–10	12	107	10–7	41
33	6–11	13	108	10–8	42
38	7–0	14	109	10–9	42
40	7–1	15	110	10–10	42
41	7–2	15	111	11–0	43
45	7–3	17	112	11–1	43
49	7–4	18	113	11–2	44
51	7–5	19	114	11–3	44
52	7–6	19	115	11–4	45
56	7–7	20	116	11–6	46
59	7–8	21	117	11–7	46
61	7–9	22	118	11–8	46
62	7–10	23	119	11–10	47
64	7–11	23	120	11–11	47
67	8–0	24	121	12–0	48
68	8–1	25	122	12–2	49
70	8–2	25	123	12–3	49
71	8–3	26	124	12–4	49
74	8–4	27	125	12–5	49
76	8–5	27	126	12–7	50
77	8–6	28	127	12–8	51
78	8–7	29	128	12–10	51
81	8–8	29	129	13–0	52
83	8–9	30	130	13–1	52
84	8–10	31	131	13–3	53
85	8–11	31	132	13–4	53
87	9–0	32	133	13–6	53
88	9–1	32	134	13–8	54
89	9–2	33	135	13–9	54
90	9–3	33	136	13–10	55
91	9–4	34	137	14–0	55
93	9–5	35	138	14–1	56
94	9–6	35	139	14–3	56
95	9–7	36	140	14–5	57
96	9–8	37	141	14–7	57
97	9–9	37	142	14–9	58
99	9–10	37	143	14–10	58
100	10–0	38	144	14–11	58

In the grade scale raw scores are transmuted into grade scores which measure achievement and which show the average grade level for the achievements of the various raw scores. The process of transmutation of raw scores into grade scores is described in detail.

In the age scale the achievement measure is the age score. This score is derived from either the age norm or the grade norm by a process described in detail which enables a teacher to make age tables, grade tables, and age-grade tables to use for any test given.

In the T scale the achievement measure is the T score, which measures a pupil in terms of the average achievement of pupils twelve years old. Detailed steps are given for finding the T score. If the T score is greater than 50 it indicates an achievement greater than the average achievement of twelve-year-old children; if it is less than 50 it indicates an achievement less than the mean achievement of twelve-year-old children.

SELECTED BIBLIOGRAPHY

FENTON, N., and WORCESTER, D. A. An Introduction to Educational Measurement, chap. iii. Ginn and Company, Boston, 1928.

GREENE, H. A., and JORGENSEN, A. N. The Use and Interpretation of Educational Tests, chap. vii. Longmans, Green & Co., New York, 1929.

GREGORY, C. A. Fundamentals of Educational Measurement, chap. v. D. Appleton and Company, New York, 1922.

MCCALL, W. A. How to Measure in Education, chap. iv, sect. 3; chap. x, pp. 256–263. The Macmillan Company, New York, 1922.

OTIS, A. S. Statistical Method in Educational Measurement, chap. xiv. World Book Company, Yonkers-on-Hudson, 1925.

CHAPTER X

MEASURES OF ABILITY

Meaning and importance of ability. Of the elements which have become important in modern education, one of the more difficult with which to deal and to interpret is that of ability. It has long been known that there were certain influences of heredity which contributed to the development of children in their schooling; but, except within wide limits, it has until recently been impossible to isolate them. One of the more important hereditary traits from the point of view of the teacher — more important than race, more important than color, and more important than sex (although this has been conceded but recently) — is the trait of brightness, because it is the degree of brightness with which an individual is endowed by inheritance that makes it possible for him to learn as he does or may. The degree of brightness practically determines what an individual may do, not necessarily (as has been argued) so much in terms of kind, but at least in terms of rate and quality.

Brightness is frequently termed "intelligence," but there is about the word "intelligence" a connotation which is somewhat undesirable. To say that an individual is lacking in intelligence is an affront. It is something more than to say that he is lacking in weight or in height, and the meaning is frequently misconstrued. Yet people are perfectly willing to concede to others in terms of ability or in terms of brightness, possibly because these terms have had less technical usage.

In reality differences in ability or brightness or intelligence, as they are known, are far less under the control of

environmental factors (if at all) than are height and weight, and are therefore less capable of being judged.

So far as the school is concerned it has always been as difficult for a teacher to judge the brightness of pupils as to judge the quality of the work that they may have done. There seems to be little correlation between facial expression and brightness, or in the appearance of an individual and brightness, or in the size or general physical characteristics of an individual and brightness. The result has been that many pupils have been misjudged — some by being thought more capable than they really were, and others by being thought to be less capable.

A pupil should not be expected to learn faster than his ability would warrant, any more than he should be expected to stretch higher than his normal reach, but the pupil who can learn quickly should be expected to do so. For the duller pupil to be pushed beyond reason can result in little but misery, whereas for the brighter pupil not to be urged to do what he can is practically sure to result in laziness, at least in school work.

The development of tests for the determination of these inherent differences in pupils has been described in Chapter III, p. 38. It has been found possible to devise tests which require of pupils certain reactions that are a measure of their inherited abilities. Certain reactions have been found to be characteristic of pupils of given ages and above, and certain reactions have also been found where the rate of accomplishment varies at different ages. These and other elements have been combined in ingenious ways to form tests which are indicative of the person's general brightness, ability, or intelligence and which give to teachers a means whereby they can judge more truly the capacities of their pupils for learning. Although known as intelligence tests, they probably measure some things dependent upon intelligence.

Tests for intelligence or ability. Tests for ability can be divided primarily into two general classes: individual tests and group tests. The individual tests were the first to be developed, as has already been pointed out, and have been brought to a high stage of perfection for use with pupils in America. Although they have a high reliability, there are two outstanding difficulties which prevent their extensive use by teachers. The first of these difficulties is the fact that individual tests must be given to pupils by one who is expert. As in all intelligence tests the directions and practice exercises are a part of the tests themselves, and in the individual tests particularly there is much room for the examiner who is inexperienced to vary the directions by voice inflection or by neglect of time and the like, so as to render the results obtained somewhat unreliable. A second difficulty in the use of individual intelligence tests by the average teacher is the amount of time consumed in giving them. Each pupil must be treated separately; and since the administration of a complete test for a single individual frequently consumes an hour, the process is too wasteful of time for the average teacher to attempt.

Attempts were made to devise tests of ability or intelligence which would remedy these two defects, and, largely as a result of the success of the tests in the United States army during the World War, technics were developed for so doing in the testing of school children.

The disadvantage of the individual intelligence test in requiring pupils to be dealt with individually, and thus wasting time, has been eliminated through constructing the test in such a way that the necessary directions can be given by the examiner to a large group at one time, and the individuals in the group can record their reactions without a personal contact with the examiner. The other main difficulty, that of the variability of the instructions having weight in

the final result, has been eliminated in great measure through constructing the test so that the directions are relatively simple to give and follow and so that the interpretation of the results is in terms of a fixed key rather than through any subjective interpretation of an oral answer.

Of group tests there are two kinds: language and non-language tests. So far as the teacher is concerned, the non-language tests are designed primarily for use in those grades where reading ability is low or absent. The large majority of pupils in the public schools can utilize the language tests. The main handicaps lie in those grades or schools where the pupils are unfamiliar with the reading, writing, and speaking of English, as is true in certain districts in our large cities almost exclusively populated by non-English-speaking foreigners.

Precautions in selecting an intelligence test. In selecting an intelligence test for use with a group of pupils the first requisite is to choose a test and a form of that test suitable for the pupils to whom it is to be given. Kindergarten and lower primary children should be given non-language intelligence tests. Occasionally reading tests can be given with success in the first and second grades, but there is usually a fair proportion of pupils in these grades to whom reading is still an accomplishment rather than a tool, so that a language-form of test is undesirable.

In the language tests most publishers have tests designed for both lower and upper grades. This is on the ground of economy, since the lower grades are able to do less than the upper grades. Tests are therefore made up as for the lower grades. They contain the same elementary material but lack extended materials designed for testing the pupils in the upper grades. Before ordering tests to be used for a class the teacher should find out which type of edition will best suit his needs. As a rule the editions of tests for the lower grades

are less expensive than those for the upper grades because of their smaller size and restricted content.

If pupils have been given an intelligence test before in the same school, and if the same intelligence test is to be given again, the teacher should endeavor to discover exactly what form of the test has been given earlier so that a different form of the same test can be ordered. The intelligence tests are usually published in several forms, so that it is not a difficult matter for a teacher to be able to utilize the same test with the same pupils, using a different form.

As a general rule a teacher is likely to have better success using a test with which he is familiar than using one which he has not seen before. There is little choice among intelligence tests, save in terms of their special application, such as the testing of certain traits and the like, which are explained by the publishers in the descriptions of the tests.

Precautions in giving an intelligence test. In giving a test the first requisite is the selection of a suitable time. A test may be given at any time during the day; but it is probably better to administer it some time during the morning session rather than during the afternoon session, as the pupils are likely to be less sluggish in the morning and less subject to distraction. The time selected should be one likely to be free from recesses, and visits from outsiders, including other teachers, principals, supervisors, and janitors. During an examination period the room should be closed to visitors except in the case of dire emergency.

As a preparation for the test the teacher should see to it that the blackboards are erased, that stray papers are removed from the desks, and that the room and desks are put in order for the work in hand. Anything that might influence the pupils during the progress of the testing, either as a distraction or as a distant help, should be removed.

As a rule one of the important factors in intelligence tests

is that of the time involved. The amount of time for practice exercises and for the body of any test or section of a test is always clearly stated in the direction booklet which accompanies the test materials. These times, to minutes and seconds, should be adhered to strictly throughout the testing period. The best way of accomplishing this is for the teacher to use a stop watch. Many schools now provide a stop watch containing both a minute and a second hand for the use of teachers in testing. If a stop watch is not available, an ordinary watch containing a second hand may be used, but great care must be taken to make sure that the time is kept accurately. A wrist watch without a second hand is inadequate for testing purposes. It is advisable whenever possible to have a helper to keep account of the time. It is good practice when a section of the test has been started to write down on paper the exact time of the start of that section, to add to it the amount of time allowed for the section to be completed, and to make note of that time. When the time is completed, it should be compared with the written memorandum.

Time started	10 hours	47 minutes	17 seconds
Time allowed	—	15 minutes	—
Time to stop	11 hours	2 minutes	17 seconds

If time is kept in this way, there will be little opportunity for error.

Before giving a test a teacher should practice reading the directions and should become familiar with the time elements that are given. He should not attempt to memorize the directions. This is wholly unnecessary and is never expected. However, he should be familiar enough with the directions to be able to read them fluently, clearly, and slowly enough for every child who is capable of understanding to do so. If the manual says that directions may be repeated, it is entirely proper to do so, but the repetition should be made before the

whole class and not just to one child. A safe rule to follow is not to give a child any aid that is not given to the entire group and is not specifically sanctioned in the direction manual. As has been stated, the directions and practice exercises are as a rule as much an element in the testing as is the arrangement of the test material or the elements which it contains.

When pupils are absent from a testing period arrangements should be made to give them a test at some later time. This should be the same test that was given to the other pupils, but in a different form.

Precautions in scoring an intelligence test. In scoring an intelligence test it is advisable for two persons to work together if possible, unless the test is scored largely by means of a key which is superimposed on the answers. When this is not the case, one person can read the answers from the key and a second person can do the scoring for five or more papers at once.

In general, the scorers should score one entire section of a group of test papers before proceeding to another section. In this way the scoring is likely to be more accurate as well as more rapid. When there are debatable answers, such as the leaving off of dollar signs, not writing in the word "cents," underlining two answers with one evidently crossed out, and the like, the manual usually gives specific instructions for scoring. When there are no specific instructions to the contrary, the scorer should take as the answer what was the evident intent of the pupil and mark the answer right or wrong on that basis. No implied intent or vague answer should be counted correct unless the manual specifically so states.

When the scoring has been completed, the entire group of papers should be re-scored. This is not only good practice but will tend to correct any errors made in the scoring process.

When a paper has been completely re-scored, it should be marked in some way to distinguish it.

The direction manual will give directions for finding the total raw scores of the tests. The usual process is to transfer the original raw scores to some single portion of the test (possibly the front page), and there to assemble the results for a final score. In many tests there are special ways of scoring certain parts of the tests, such as subtracting the number of wrong answers from the number of right ones, multiplying the total number of right answers by 2, and the like. Such directions should be scrupulously followed; otherwise the final results will be useless.

As a final operation in scoring the test papers a check should be made of the important data required to make the test results useful, particularly if they are to be used by any one other than the room teacher. Such data as the date on which the test was taken, the name of the pupil, his grade, and his age or birthday and year of birth are all of the greatest importance, and each item should be checked for correctness either by someone who is familiar with the data, or against a teacher's record. When the scoring has been completed, the results are ready for assembling and using.

Measures of ability. When the raw scores on intelligence tests have been obtained, it is possible to transmute them into various forms of measures of ability for the purpose of giving them meaning and usefulness. These measures of ability are in the grade scale, the age scale, and the T scale.

Measures of ability in the grade scale. The measure of ability in the grade scale is the grade score in intelligence (G_I). Ordinarily this is not so good a measure for general use as either of the measures in the other scales unless it is derived from them. If the grade score in intelligence is derived from grade norms for the test, it is likely to be less adequately adapted to a school system than if it is derived from age

norms or from T scores, because, as has already been stated, there is less variation in age achievements of pupils than there is in their grade levels.

1. *Deriving G_i from grade norms.* When the grade score in intelligence (G_i) is to be derived from grade norms, the procedure is as follows:

STEP 1. *Determining the grade norms of the test*

The grade norms for the test are published in the test manual and should be listed.

STEP 2. *Preparing a transmutation sheet*

The transmutation sheet should be prepared just as when raw scores are to be transmuted into grade scores in achievement. The process has been described in Chapter IX, pp. 203–211.

STEP 3. *Placing the norm positions on the transmutation sheet*

The norm positions are the places on the transmutation sheet where the lines indicating given grade scores cross the lines indicating given raw scores. In making these placements of norm positions on this sheet, the time of year that the norms have as their point of reference should be considered.

STEP 4. *Drawing the norm line on the transmutation sheet*

This process has been described on pages 208–209, and it should be remembered that the line so drawn should be a smoothed line and not just a series of straight lines connecting the norm points.

STEP 5. *Constructing a raw-score-G_i table from the transmutation sheet*

By reading on the transmutation sheet where each raw score line, norm line, and grade line intersect, and by transferring the grade scores indicated to the sheet containing all

the possible raw scores, a raw-score-G_i table will be constructed which will enable a teacher to transmute raw scores on intelligence tests directly into grade scores in intelligence. As a rule other scores will be desired, such as mental-age scores or T scores, and the table should later be enlarged to include these.

2. *Deriving G_i from age norms.* Practically every intelligence test places greater stress on age norms than upon grade norms, and thus the age norms for every possible raw score in the test are as a rule carefully worked out. In this way it is possible to transmute these age norms into grade scores in intelligence applicable to any given school system in the following manner:

STEP 1. *Determining the average entering age of pupils in the first grade*

This can be done by consulting Table XXVIII A when the age norms are given in years and months, or Table XXVIII B when they are given in months only (see page 227).

STEP 2. *Determining G_i for each age score*

By means of Table XXVIII A or Table XXVIII B, the general table for the transmutation of age to grade scores or of grade to age scores, it is possible, by using the proper column of ages as determined by Step 1, to transmute each age score directly from the table given in the test manual into a G_i for use in the school system. The G_i scores thus obtained will be likely to be true for the system and will have greater meaning for the teachers.

3. *Deriving G_i from T scores.* When the T scores in intelligence have been found, as will be outlined in a later section, it is possible to utilize them for determining grade scores in intelligence without the necessity of referring to grade norms or age norms. If the T scores have been obtained through the use of Table XXXI, p. 243, the operation is

simple. It involves merely the looking up on that table of the T score for any individual, reading across to the column of grade scores next to it, and thus reading directly the grade scores in intelligence.

4. *Deriving G_i from B_i scores.* When rate scores in intelligence[1] have been obtained, it is possible to transmute them into standard grade scores in intelligence by means of a table constructed by Dr. McCall. As will be pointed out in a later section this table (obtained at nominal cost through Teachers College Bureau of Publications, Columbia University, New York City) is an essential tool in the hands of anyone who proposes to use the T-G-B-F-scale system, under which B scores are used. One portion of the booklet of tables is devoted to tables for the transmutation of B scores to grade scores. It is merely necessary to locate at the left of the table the pupil's age and at the top his B score. Under the latter, and on a level with the former, is a figure which represents the grade score. If the B score is a B_i, the grade score found will be a G_i.

Measures of ability in the age scale. The measure of ability in the age scale is the mental age, abbreviated as MA.

The mental age. As in the achievement measures in the age scale, where the reading age indicates the extent of achievement in reading, the arithmetic age indicates the extent of achievement in arithmetic, and so on, and where each is in terms of the average achievement of pupils of that age, so in the measures of ability the mental age indicates the relative ability of pupils in terms of the average ability of pupils of a like age. It thus becomes possible, when the achievement in a given intelligence test has been measured for a sufficient number of pupils, to determine the raw score which is comparable to the achievement of pupils at any given age. As has been stated, nearly every intelligence test

[1] See page 272.

TABLE XXXIV. AGE NORMS FOR THE NATIONAL INTELLIGENCE TEST[1]

SCORE	N.I.T. MA		SCORE	N.I.T. MA	
Scale A	Years	Months	Scale A	Years	Months
46	8	6	91	11	5
47	8	7	92	11	5
48	8	7	93	11	6
49	8	8	94	11	6
50	8	9	95	11	7
51	8	10	96	11	8
52	8	11	97	11	9
53	8	11	98	11	10
54	9	0	99	11	10
55	9	1	100	11	11
56	9	2	101	12	0
57	9	3	102	12	0
58	9	4	103	12	1
59	9	4	104	12	2
60	9	5	105	12	2
61	9	6	106	12	3
62	9	7	107	12	4
63	9	8	108	12	5
64	9	9	109	12	6
65	9	10	110	12	7
66	9	10	111	12	7
67	9	11	112	12	8
68	9	11	113	12	9
69	10	0	114	12	10
70	10	1	115	12	11
71	10	2	116	13	0
72	10	3	117	13	0
73	10	4	118	13	1
74	10	5	119	13	2
75	10	6	120	13	3
76	10	6	121	13	4
77	10	7	122	13	4
78	10	8	123	13	5
79	10	8	124	13	6
80	10	9	125	13	7
81	10	10	126	13	9
82	10	11	127	13	10
83	10	11	128	13	11
84	11	0	129	14	0
85	11	0	130	14	1
86	11	1	131	14	2
87	11	2	132	14	4
88	11	2	133	14	6
89	11	3	134	14	7
90	11	4	135	15	1
			136	15	6

[1] By permission of the World Book Company.

places importance on the mental age, and therefore there has been published for each a list of age norms for the entire range of possible raw scores. This makes it unnecessary to construct a raw-score-age table, and to turn any raw score into a mental age it is necessary merely to read the mental ages directly from the norm table. Table XXXIV on page 259 shows the age-norm table of the National Intelligence Test as of 1922. This table was constructed as a result of the testing of 32,372 children.

TABLE XXXV. RAW SCORES IN THE NATIONAL INTELLIGENCE TEST FOR FIVE PUPILS

PUPIL	SCORE
A	76
B	117
C	97
D	62
E	103

The procedure for transposing the scores into mental ages is as follows: Suppose a group of five pupils has made the raw scores indicated in Table XXXV.

Pupil A received a raw score of 76. In Table XXXIV the age for this score is found to be 10–6, which can then be assigned to that score in Table XXXV.

Pupil B received a raw score of 117. Table XXXIV shows this to be equivalent to a mental age of 13–0, which can then be assigned to that pupil.

In the same way, the raw scores of the three other pupils can be transmuted into mental ages with the result as shown in Table XXXVI.

TABLE XXXVI. MENTAL AGES FOR FIVE PUPILS IN TABLE XXXV AS FOUND FROM TABLE XXXIV (N.I.T.)

PUPIL	RAW SCORE	MENTAL AGE
A	76	10–6
B	117	13–0
C	97	11–9
D	62	9–7
E	103	12–1

Measures of ability in the T scale. In the T scale the measure of ability in use is the T score in intelligence, known as T_i.

The T score in intelligence (T_i). The T score in intelligence indicates the relationship of an individual to the dis-

tribution of twelve-year-old children in intelligence. The measure is on the same scale as that of the T score in the achievement measures, a scale of from 0 to 100 T, where 50 T represents the average T rating of intelligence of twelve-year-old children (12–6). The steps in finding the T_i are as follows:

STEP 1. *Making a T table for the raw scores of the test*

When a T table for the raw scores of a test is not available, it can easily be made up from the table of raw scores and mental ages that is furnished with the test manual. This is done in the same manner as has been described for the finding of achievement T scores.[1] By consulting Table XXXI, p. 243, the T scores for the various ages given in the test manual can be made up into a T table for use with the raw scores. Table XXXVII on page 262 shows such a T table as made up for the raw scores in the National Intelligence Test. It will be noted that this table varies slightly from that given in Table XXXI. This is because it is based upon this test alone, and it is therefore somewhat more accurate for this test.

STEP 2. *Transmuting the raw scores for pupils into T scores in intelligence*

This step merely involves the transmutation of the raw scores as found on the papers of the pupils into T scores in intelligence by means of the T table, which has been constructed.

For the pupils listed in Table XXXV, Pupil A has a raw score of 76. By consulting Table XXXVII, the raw-score-T table, a raw score of 76 is equivalent to a T score of 40. Pupil B has a raw score of 117. His T score, according to Table XXXVII, is 52. The T scores found in this way are as listed in Table XXXVIII on page 263.

[1] See page 245.

TABLE XXXVII. T-G TABLE FOR NATIONAL INTELLIGENCE TEST, SCALE A[1]

RAW SCORE	T	G	RAW SCORE	T	G
0–2	20	0.7	140–142	60	8.6
3–4	21	0.9	143–144	61	8.8
5–7	22	1.1	145–146	62	9.0
8–10	23	1.3	147–148	63	9.2
11–13	24	1.5	149–151	64	9.3
14–15	25	1.7	152–153	65	9.5
16–18	26	1.9	154–155	66	9.7
19–21	27	2.1	156–157	67	9.9
22–25	28	2.3	158–159	68	10.1
26–29	29	2.5	160–161	69	10.3
30–33	30	2.6	162–163	70	10.5
34–37	31	2.8	164	71	10.7
38–42	32	3.0	165–166	72	10.9
43–47	33	3.2	167	73	11.1
48–51	34	3.4	168	74	11.3
52–56	35	3.6	169	75	11.5
57–61	36	3.8	170	76	11.7
62–65	37	4.0	171–172	77	11.9
66–69	38	4.2	173	78	12.1
70–72	39	4.4	174	79	12.3
73–76	40	4.6	175	80	12.5
77–79	41	4.8	176	81	12.7
80–83	42	5.0	177	82	12.9
84–87	43	5.2	178	83	13.1
88–91	44	5.4	179	84	13.3
92–94	45	5.6	180	85	13.5
95–98	46	5.8	181	86	13.7
99–102	47	6.0	182	87	13.9
103–105	48	6.2	183	88	14.1
106–108	49	6.4	184	89	14.3
109–110	50	6.6	185	90	14.5
111–113	51	6.8	186	91	14.7
114–117	52	7.0	187	92	14.9
118–120	53	7.2	188	93	15.1
121–123	54	7.4	189	94	15.3
124–127	55	7.6	190	95	15.5
128–131	56	7.8	191	96	15.7
132–133	57	8.0	192	98	16.1
134–136	58	8.2	193	99	16.3
137–139	59	8.4	194	100	16.5
			195	101	16.7
			196	103	17.0

[1] By permission of Teachers College Bureau of Publications.

Keeping mental ages up to date. It is not necessary to give a new intelligence or ability test each time it is desired to use the mental ages or B scores from them. The value of these scores, provided they are derived from good tests carefully given, is relatively permanent. It should not be necessary to derive them oftener than once in two or three years, but to be of value they must be readjusted constantly to make them useful.

TABLE XXXVIII. T SCORES OF PUPILS OF TABLE XXXV

PUPIL	RAW SCORE	T SCORE
A	76	40
B	117	52
C	97	46
D	62	37
E	103	48

A record for each pupil should be kept of the exact date upon which an intelligence test was given. For this particular test let it be assumed that the record was originally made in terms of a mental age, which is also recorded. To correct this mental age for any later date, it is necessary first to find the intelligence quotient, the process for which is described in Chapter XI. The intelligence quotient should also become a part of this permanent record.

TABLE XXXIX. SHOWING MENTAL AGES AND INTELLIGENCE QUOTIENTS FOR PUPILS IN TABLE XXXVI

PUPIL	CA	MA	IQ
A	10–6	10–6	100
B	14–2	13–0	91
C	10–7	11–9	111
D	8–9	9–7	109
E	12–3	12–1	98

Whenever a new mental age is desired, it is merely necessary to multiply the new chronological age by the intelligence quotient, which will result in the new mental age for the pupil concerned.

In the case of the pupils in Table XXXIX let it be assumed that seven months after the giving of the intelligence test a new mental age was desired for each child.

Pupil A, who had a chronological age of 10–6, or 126 months, now has a chronological age of 11–1, or 133 months.

Since his IQ is relatively permanent, to find the new mental age the new chronological age should be multiplied by the IQ. In these cases the IQ should be pointed off two places in order to restore the decimal points, which will be found to have been removed for the sake of convenience.[1]

$$CA \times \frac{IQ}{100} = MA.$$

$$133 \times 1.00 = 133, \text{ or } 11\text{–}1,$$

In the same way, Pupil B with an original CA of 14–2, or 170 months, now has a CA of 14–9, or 177 months. His IQ is 91, so the new mental age is found as follows:

$$CA \times \frac{IQ}{100} = MA.$$

$$177 \times 0.91 = 161, \text{ or } 13\text{–}5.$$

When a similar procedure has been carried out for the remaining three pupils of this group, the results, giving the new mental ages in terms of the new chronological ages, are as shown in Table XL.

TABLE XL. REVISED MENTAL AGES WITH INCREASE IN CHRONOLOGICAL AGE OF SEVEN MONTHS

PUPIL	CA AT TEST	NEW CA	INITIAL MA	FINAL MA	IQ
A	10–6	11–1	10–6	11–1	100
B	14–2	14–9	13–0	13–5	91
C	10–7	11–2	11–9	12–5	111
D	8–9	9–4	9–7	10–1	109
E	12–3	12–10	12–1	12–7	98

Chapter summary. Measures of ability constitute one of the newer additions to the science of measurement. They measure the extent of the hereditary traits which are responsible for learning, and are therefore of great value to the

[1] See page 270.

teacher in aiding him to understand the successes or failures of pupils in their dealing with school learnings.

Measures of ability are obtained through the giving of intelligence tests to pupils. The tests which are given should be selected with care for their fitness for the group and for the degree of the probable abilities of the group.

In giving an intelligence test a suitable time, free from distractions, should be selected. Pupils should be supplied with necessary materials, and any possible helps in completing the test which are present in the schoolroom should be removed. Since time is usually an important factor, the timing of the test or of the giving of various portions of it should be as accurate as possible. The teacher should either use a stop watch or be assisted by a timekeeper using a watch with a second hand.

In scoring an intelligence test two persons can usually work more effectively than one. The scoring should be done carefully in terms of the keys provided in the test manual, and the scores should be re-checked to eliminate error.

The measure of ability in the grade scale is the grade score in intelligence (G_i). This can be obtained from grade norms furnished with the test, from T scores derived from the raw scores of the test, or from B_i scores derived from the T scores. Detailed directions for transmuting raw scores into grade scores in any one of the four ways are included in the discussion.

The measure of ability in the age scale is the mental age (MA). The mental age is analogous to the achievement ages in the same scale, showing a level of ability rather than a level of achievement. The mental age is found directly from the raw scores of a test from a table in the manual of directions.

The measures of ability in the T scale are the T score in intelligence (T_i) and the B score in intelligence (B_i). The T_i

score shows the relationship of the ability of a pupil as compared with twelve-year-old pupils (12–6) on the T scale.

The mental age must be corrected continually by multiplying new chronological ages by the IQ's, in order to keep them useful in further operations and interpretations. This can be done; for if an intelligence test is carefully given and carefully checked, it is a relatively permanent measure.

SELECTED BIBLIOGRAPHY

BUCKINGHAM, B. R. Research for Teachers, chap. iii. Silver, Burdett and Company, New York, 1926.

FENTON, N., and WORCESTER, D. A. An Introduction to Educational Measurements, chap. ii. Ginn and Company, Boston, 1928.

MCCALL, W. A. How to Measure in Education, pp. 21–22, 38–40. The Macmillan Company, New York, 1922.

PINTNER, R. Intelligence Testing, p. 73. Henry Holt and Company, New York, 1923.

TERMAN, L. M. The Measurement of Intelligence. Houghton Mifflin Company, Boston, 1916.

THORNDIKE, E. L. The Measurement of Intelligence. Teachers College Bureau of Publications, New York, 1927.

TRABUE, M. R. Measuring Results in Education, chap. xvii. American Book Company, New York, 1924.

CHAPTER XI

MEASURES OF RATE

Significance of rate in learning. Chapters IX–X have considered the measurement of achievement and the measurement of ability. Achievement is not merely what a pupil has accomplished, but rather what he has accomplished as a result of all the factors of his nature and environment which have influenced him. One of the factors of his nature is his ability, and, as has been implied and as will be shown more fully here, his ability is an important factor in determining the extent of achievement. However, to understand the relationship of ability and achievement it is necessary to understand a pupil's rate of learning. As will be seen, there are various aspects of this rate of learning, each of which bears a relation to achievement or ability. There is, first, the potential rate of learning. This is that rate of achieving which a pupil should possess if he were to realize to the fullest the promise of his ability. There is also the rate of learning which has resulted in the achievement of a pupil and which has been affected by many other factors besides mere ability. There is, finally, a probable future rate of learning which can be estimated as desirable for a pupil.

Rate of learning is an important element of knowledge for a teacher to determine for each pupil, since it will serve to correct false impressions with respect to a pupil, will enable a teacher to judge more accurately the quality of work which he produces, and will allow a better judgment as to what can reasonably be expected of him. In dealing with a class of pupils a teacher can adjust the work to meet the needs of the

class if the rate of learning of the group is known. If the rate of learning has been determined, teachers themselves are less likely to be misjudged. Not only is a pupil, if his rate of learning is below normal, subject to misjudgment on the part of teachers who urge him to exceed his normal development, but teachers of pupils whose rate of learning is below normal development are compared frequently with teachers whose pupils are of normal or above normal in rate of learning, to the disadvantage of the former. In promoting pupils, in grading them, in judging the quality of their work, and in dealing with them in their daily work, to know the rates of learning will enable a teacher to deal more fairly, to promote more wisely, and to judge them and their products more clearly.

Measures of potential rate in learning. If other factors, such as interest, laziness, difficulties in subject matter, absence, and the like did not enter to affect the achievement of pupils, it would not be necessary for a teacher to determine the potential rate of learning, for their actual progress in learning would be sufficient. However, such factors do enter to increase or decrease the normal rate of learning. Frequently pupils are stimulated in certain fields so that their rate of learning is above normal, but more often they are retarded. It is then important that a teacher should know what is a normal rate of learning for any pupil regardless of stimulation or retardation, in order that comparisons with the actual achievement can be made and corrections instituted.

The potential rate of a pupil's learning is the rate at which he should learn if his full measure of ability is utilized. This rate is expressed in terms of some comparison to the rate of learning of normal pupils of the same age as the pupil in question. This comparison shows that a pupil is more rapid in his potential rate of learning than the normal

pupil of the same chronological age, that he learns more slowly, or that he is equal in rate of learning to pupils of his own age.

In the *age scale* the measure of the potential rate of learning is the intelligence quotient; in the *T scale* the measure of potential rate of learning is the B score in intelligence.

Determination of the intelligence quotient. The intelligence quotient, usually known as the IQ, is a relationship between the ability of a pupil and the ability of pupils of his own chronological age. It can be found in terms of the ability ratings in the age scale, which have been discussed previously.

STEP 1. *Determining the mental age*

As has already been stated, the mental age is determined from an intelligence test.[1] By means of a transmutation table of raw scores to age scores the raw scores for pupils from an intelligence test can be transmuted into age scores, which are called mental ages.

STEP 2. *Determining the chronological age*

The determination of chronological age has been discussed in Chapter V and should be determined for the same time as that at which the intelligence was given.

STEP 3. *Reducing mental and chronological ages to months*

The mental ages as found and the chronological ages of the pupils should be transposed into ages expressed in months. This can be done by means of Table XLI. For the five pupils for whom mental ages were found in Table XXXVI, p. 260, the transposition of mental ages into months is shown in the second and third columns of Table XLII, and the transposition of chronological ages of the same pupils is shown by means of the fourth and fifth columns.

[1] See page 258.

TABLE XLI. FOR REDUCING AGES EXPRESSED IN YEARS AND MONTHS TO AGES EXPRESSED IN MONTHS ONLY

YEARS	MONTHS											
	0	1	2	3	4	5	6	7	8	9	10	11
5	60	61	62	63	64	65	66	67	68	69	70	71
6	72	73	74	75	76	77	78	79	80	81	82	83
7	84	85	86	87	88	89	90	91	92	93	94	95
8	96	97	98	99	100	101	102	103	104	105	106	107
9	108	109	110	111	112	113	114	115	116	117	118	119
10	120	121	122	123	124	125	126	127	128	129	130	131
11	132	133	134	135	136	137	138	139	140	141	142	143
12	144	145	146	147	148	149	150	151	152	153	154	155
13	156	157	158	159	160	161	162	163	164	165	166	167
14	168	169	170	171	172	173	174	175	176	177	178	179
15	180	181	182	183	184	185	186	187	188	189	190	191
16	192	193	194	195	196	197	198	199	200	201	202	203
17	204	205	206	207	208	209	210	211	212	213	214	215

TABLE XLII. REDUCTION OF AGES, MENTAL AND CHRONOLOGICAL, OF PUPILS LISTED IN TABLE XXXVI

PUPIL	MA (Years)	MA (Months)	CA (Years)	CA (Months)
A	10–6	126	10–6	126
B	13–0	156	14–2	170
C	11–9	141	10–7	127
D	9–7	115	8–9	105
E	12–1	145	12–3	147

STEP 4. *Dividing the mental age by the chronological age*

The intelligence quotient is found from the mental and chronological ages, when these have been converted into months, by means of the simple formula given below. It will be noted that the quotient has been multiplied by 100. This is a common practice and is done in order to remove the decimal points, but the true quotient should be understood to be some number greater than, equal to, or less than 1.

$$(100 \times) \frac{MA \text{ (mental age)}}{CA \text{ (chronological age)}} = IQ \text{ (intelligence quotient).}$$

For Pupil A of Table XLII the mental age in months is 126. His chronological age, likewise reduced to months, is 126. The substitution in the foregoing formula becomes

$$(100 \times) \frac{126\ (MA)}{126\ (CA)} = 100\ (IQ).$$

For Pupil B, whose mental age is 156 months and whose chronological age is 170 months, the substitution is as follows:

$$(100 \times) \frac{156\ (MA)}{170\ (CA)} = 91\ (IQ).$$

When the calculations have been made for the remaining pupils, the results are as shown in Table XLIII.

Interpretation of intelligence quotients. The intelligence quotient shows the extent to which an individual varies in ability from normal individuals of his own age, and is to that extent a measure of potential rate in learning. Owing to the fact that the present measures of intellect or ability may not be perfect or may not be equally adapted to the measurement of all individuals, certain investigators have found variations in the intelligence quotients of the same individual as measured by different tests at different times. Pintner points out that it is possible that the rate of growth of intelligence may be somewhat irregular in an individual in some cases, but that the usual growth of ability in individuals is regular until a maximum is reached at between fourteen and sixteen years of age.[1] During this period the intelligence quotient, since it is a measure of rate, is in most

TABLE XLIII. THE INTELLIGENCE QUOTIENTS OF FIVE PUPILS OF TABLE XLII

PUPIL	MA	CA	IQ
A	126	126	100
B	156	170	91
C	141	127	111
D	115	105	109
E	145	147	98

[1] Rudolph Pintner, *Intelligence Testing*, p. 70. Henry Holt and Company, New York, 1923.

cases fairly stable; therefore it can be used repeatedly from year to year, as will be shown later. (See also page 263).

The meaning of the intelligence quotient in terms of relative ability has been given roughly by grouping intelligence quotients and giving the characteristic of the individuals whose intelligence falls within any group. Thus Pintner gives the following (Table XLIV) as the meaning of intelligence quotients. It should be understood that there is no abrupt change or difference between the individuals close together in intelligence quotients but in different groups.

TABLE XLIV. INTERPRETATION OF INTELLIGENCE QUOTIENTS

(After Pintner, p. 77)

INTELLIGENCE QUOTIENTS	CLASSIFICATION
0–69	Feeble-minded
70–79	Border-line
80–89	Backward
90–109	Normal
110–119	Bright
120–129	Very bright
130 and upward	Very superior

As can be seen from this table an intelligence quotient above 100 is above normal, one below 100 is below normal, and quotients of 100 represent normality.

Determination of the B score in intelligence. In the T scale the B score in intelligence, commonly called the B_i, is analogous to the intelligence quotient in the age scale.

It will be remembered that the T score in intelligence is a score, regardless of the age of the pupil, which showed the relationship of the pupil in intelligence to the distribution of twelve-year-old children. It is not expected that a normal child ten years old should have as high a T score in intelligence[1] as one twelve years old or that a normal child of four-

[1] See page 261.

teen years should have as low a T score in intelligence as one twelve years old. Ten-year-old pupils have a range of intelligence similar to that of twelve-year-old pupils, as do also those who are fourteen years old. This situation is shown graphically in Fig. 78. Here the distribution of twelve-year-old pupils is shown extending from 0 T to 100 T. A similar distribution of ten-year-old children is also shown, as well as that of fourteen-year-olds. Three scales are shown in the figure: first, the T scale, which shows the T units for the

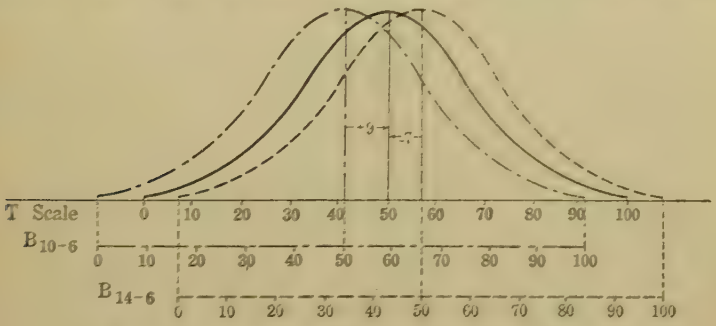


FIG. 78. Derivation of B scales from the T scale

children with an average age of 12–6; immediately below the T scale a scale made up of units of the same size as T units for the children with an average age of 10–6; and below this another scale, also made up of T-size units, for children with an average age of 14–6.

The mean of the T scale is 50 T; the mean of the scale for 10–6 children is 41 T, as indicated by the line drawn in the middle of the curve for 10–6; the mean for the children aged 14–6 is 57 T, as indicated by the line drawn in the middle of that curve. This means that the normal ability of pupils aged 10–6 is 9 T units less than the normal ability of pupils aged 12–6, and that the normal ability of pupils aged 14–6 is 7 T units more than that of pupils aged 12–6. To express the

units on these new scales in terms comparable with the units on the T scale, the new scales are called B scales (*B* indicating "brightness"), and the mean for the children in any of the brightness scales is 50 B. Thus 41 T on the T scale becomes 50 B_{10-6}, on the B scale for children aged 10–6, and 57 T on the T scale becomes 50 B_{14-6}.

As will be shown later, B scores in other fields than intelligence are derived in the same way as in the illustration given on page 273.

It can now readily be seen that any T score can be changed into a B score if the age of the pupil is known and if the difference in T units between the mean of the T scale and the mean of the scale for the age of the pupil is known. The steps for the transmutation are as follows:

TABLE XLV. T_i SCORES AND CHRONOLOGICAL AGES OF PUPILS OF TABLE XXXV

PUPIL	CA	T_i
A	10–6	40
B	14–2	52
C	10–7	46
D	8–9	37
E	122–3	48

STEP 1. *Determining the chronological age*

The chronological age of the pupil is determined as described on page 95. This can be in terms of years and months, and not necessarily in months, as is required in the determination of the IQ. The table which uses these ages is given in terms of years and months.

STEP 2. *Determining the T score in intelligence* (T_i)

The T score in intelligence is determined by transmuting the raw scores on an intelligence test by means of a raw-score-T table, into T scores in intelligence. The T_i scores for the five pupils considered in this chapter were found as listed in Table XXXVIII, p. 263. Their chronological ages are as found in Table XLII, p. 270. These are given in Table XLV shown above.

TABLE XLVI. B-CORRECTION TABLE. TABLE 1 OF MCCALL'S TABLES FOR THE T-G-B-F-SCALE SYSTEM

AGE IN YEARS AND MONTHS	BCR	AGE IN YEARS AND MONTHS	BCR	AGE IN YEARS AND MONTHS	BCR	AGE IN YEARS AND MONTHS	BCR
3–0	80	7–0	36	11–0	7	15–0	−9
3–1	79	7–1	35	11–1	6	15–1	−9
3–2	78	7–2	34	11–2	6	15–2	−9
3–3	78	7–3	33	11–3	5	15–3	−10
3–4	77	7–4	32	11–4	5	15–4	−10
3–5	76	7–5	31	11–5	4	15–5	−10
3–6	75	7–6	31	11–6	4	15–6	−10
3–7	74	7–7	30	11–7	4	15–7	−11
3–8	73	7–8	29	11–8	4	15–8	−11
3–9	73	7–9	28	11–9	3	15–9	−11
3–10	72	7–10	27	11–10	3	15–10	−11
3–11	71	7–11	26	11–11	2	15–11	−12
4–0	70	8–0	26	12–0	2	16–0	−12
4–1	69	8–1	25	12–1	1	16–1	−12
4–2	68	8–2	25	12–2	1	16–2	−12
4–3	68	8–3	24	12–3	1	16–3	−13
4–4	67	8–4	23	12–4	1	16–4	−13
4–5	66	8–5	22	12–5	1	16–5	−13
4–6	65	8–6	22	12–6	0	16–6	−13
4–7	64	8–7	21	12–7	−1	16–7	−14
4–8	63	8–8	21	12–8	−1	16–8	−14
4–9	63	8–9	20	12–9	−1	16–9	−14
4–10	62	8–10	19	12–10	−1	16–10	−14
4–11	61	8–11	18	12–11	−2	16–11	−14
5–0	60	9–0	18	13–0	−2	17–0	−14
5–1	59	9–1	17	13–1	−3	17–1	−15
5–2	58	9–2	17	13–2	−3	17–2	−15
5–3	58	9–3	16	13–3	−3	17–3	−15
5–4	57	9–4	16	13–4	−3	17–4	−15
5–5	56	9–5	15	13–5	−4	17–5	−16
5–6	55	9–6	15	13–6	−4	17–6	−16
5–7	54	9–7	14	13–7	−4	17–7	−16
5–8	53	9–8	14	13–8	−4	17–8	−16
5–9	53	9–9	13	13–9	−5	17–9	−16
5–10	52	9–10	13	13–10	−5	17–10	−17
5–11	51	9–11	12	13–11	−5	17–11	−17
6–0	50	10–0	12	14–0	−5	18–0	−17
6–1	48	10–1	11	14–1	−6	18–1	−18
6–2	47	10–2	11	14–2	−6		
6–3	45	10–3	10	14–3	−7		
6–4	44	10–4	10	14–4	−7		
6–5	43	10–5	9	14–5	−7		
6–6	42	10–6	9	14–6	−7		
6–7	41	10–7	8	14–7	−8		
6–8	40	10–8	8	14–8	−8		
6–9	39	10–9	8	14–9	−8		
6–10	38	10–10	8	14–10	−8		
6–11	37	10–11	7	14–11	−9		

STEP 3. *Determining the B correction (BCr) for each pupil*

The B correction is a number representing T units which added to or subtracted from the T score will reduce the T score of the pupil on the T scale to the B score on the B scale for his age.

According to Fig. 78, p. 273, if a pupil whose age is 10–6 has a T_i score of 41, the B correction would be + 9, and this added to his T score would result in his B score, which in this case is 50. If a pupil of the age 14–6 has a T score of 57, and if the B correction for his age is −7, his B score is 50 B.

TABLE XLVII. B CORRECTIONS FOR PUPILS OF TABLE XLV

PUPIL	CA	BCR
A	10–6	9
B	14–2	−6
C	10–7	8
D	8–9	20
E	12–3	1

The B correction for any pupil may be found from a table made by Dr. McCall and contained in his pamphlet of tables for the T-G-B-F-scale system. This B-correction table is reproduced as Table XLVI on page 275.

To find the B correction for any pupil by means of this table, refer to the column of ages there given for the age of the pupil in years and months. In the next column to the right opposite this age will be found the amount of the B correction for that age. This correction is a plus quantity for ages less than 12–6 and a minus quantity for ages greater than 12–6. The B score for the age 12–6 is the same as the T score, and the BCr is 0. The T scale is really, therefore, the B_{12-6} scale.

For the data contained in Table XLV, the B corrections found from Table XLVI are given in Table XLVII.

STEP 4. *Adding the B correction to, or subtracting it from, the T score*

The formula for finding the B score is as follows:

T score + BCr (regarding signs) = B score.

Applying this formula to the T score of Pupil A, the substitution becomes

$$40(T) + 9(BCr) = 49 \text{ (B score)}.$$

When the substitution has been made for the other pupils, the final B scores are as shown in Table XLVIII.

TABLE XLVIII. DERIVATION OF B_i SCORES FROM T_i SCORES FOR PUPILS OF TABLE XLVII

PUPIL	T_i	BCr	B_i
A	40	+9	49
B	52	−6	46
C	46	+8	54
D	37	+20	57
E	48	+1	49

Interpretation of the B score in intelligence. As has been said, the B_i score is a measure of potential rate of learning. If it is high above 50, the rate of learning is rapid; if it is 50, the rate of learning is equivalent to that of a normal child; if it is below 50, the rate of learning is slow. B scores can be somewhat better interpreted according to Table XLIX or Table L, which are descriptive.

Table XLIX shows the percentage of pupils of any given age who in rate of learning exceed pupils with the B scores as given in the first column. To read this table one can say that 99 per cent of pupils of the same chronological age as the pupil who receives a B score of 25 exceed him in rate of learning, that 69 per cent of pupils of the same chronological age exceed in rate of learning the pupil who receives a B score of 45, and so on for other B scores.

TABLE XLIX. INTERPRETATION OF B SCORES

B SCORE	PERCENTAGE OF PUPILS OF SAME AGE WHO SURPASS THIS SCORE
25	99
30	98
35	93
40	84
45	69
50	50
55	31
60	16
65	7
70	2
75	1
80	0.1

Table L shows the characterization which may be given to B_i scores of different sizes. It is read by saying that a pupil who receives a B_i score below 23 is exceptionally inferior in intelligence to pupils of his own chronological age, that those whose scores range from 23 to 31 are very inferior, and so on.

Dr. McCall has been able to construct a transmutation table of B_i scores and intelligence quotients by means of which any B_i score can be transmuted into an IQ, or vice versa. This table, extended somewhat from that given by Dr. McCall in his pamphlet of T-G-B-F tables, is given in Table LI. It can also be used to transmute any quotient into any B score, or vice versa.

TABLE L. CHARACTERISTICS OF B-SCORE GROUPINGS

NOTE. These interpretations may be used with equal value for B_e scores and B_i scores.

B SCORES	CHARACTERISTICS
10–22	Exceptionally inferior
23–31	Very inferior
32–38	Inferior
39–46	Low average
47–53	Average
54–61	High average
62–68	Superior
69–74	Very superior
75–80	Exceptionally superior

Measures of previous rate in learning. The measure of a pupil's potential rate of learning is a measure of how rapidly he ought to learn; the measure of his previous rate of learning is the measure of how rapidly he actually has learned. If all conditions both within and without the pupil were normal or favorable, the previous rate of learning would approximate the rate limit as set by his native ability; but many other factors enter, as has been said, to retard the learning and (occasionally) to accelerate it. It is important for the teacher to determine how greatly these conditions have affected the rate of learning of pupils in order to be able to judge pupils wisely.

In the *age scale* the measure of previous rate of learning is the achievement quotient; in the *T scale* it is the B score in education.

TABLE LI.[1] FOR TRANSMUTING ANY B SCORE INTO ANY QUOTIENT

NOTE. When transmuting quotients to B scores and B scores to quotients, use figures with asterisks as preferable choices.

B SCORE	QUOTIENT	B SCORE	QUOTIENT	B SCORE	QUOTIENT
10	55	*43	91	71	130
11	56	43	92	*72	131
12	*56	44	93	72	132
13	57	45	94	73	133
14	58	46	95	*74	134
15	59	*47	96	74	135
16	60	47	97	*75	136
17	61	48	98	75	137
18	*61	49	99	*76	138
19	62	50	100	76	139
20	63	51	101	77	140
21	64	*52	102	*78	141
22	65	52	103	78	142
23	*65	53	104	*79	143
24	66	54	105	79	144
25	67	55	106	*80	145
26	68	*56	107	80	146
*27	69	56	108	*81	147
27	70	57	109	81	148
28	71	*58	110	82	149
29	72	58	111	*83	150
30	73	59	112	83	151
*31	74	60	113	*84	152
31	75	*61	114	84	153
32	76	61	115	*85	154
33	77	62	116	85	155
*34	78	63	117	*86	156
34	79	*64	118	86	157
35	80	64	119	*87	158
36	81	*65	120	87	159
*37	82	65	121	*88	160
37	83	66	122	88	161
*38	84	*67	123	89	162
38	85	67	124		
39	86	68	125		
40	87	*69	126		
*41	88	69	127		
41	89	*70	128		
42	90	70	129		

[1] Adapted from a table in McCall's *Tables for the T-G-B-F System.*

Determination of the achievement quotient.[1] The achievement quotient, usually known as the AQ, is a relation between a pupil's achievement and that of normal pupils of his own age. It therefore shows how well he has done in comparison with other pupils like himself in age and is a measure of his previous rate of learning.

The achievement quotient may be expressed as a quotient in any form of subject matter; when so expressed, it is named in terms of the subject tested. A measure of previous rate of learning in reading is called a reading quotient, designated as ReQ; a measure of previous rate of learning in arithmetic is called an arithmetic quotient, designated as ArQ; and so on. From a number of such quotients, or from the combined data resulting in such quotients, a composite quotient can be obtained which is called an educational quotient (EQ). These quotients are found as follows:

TABLE LII. READING AGES OF FIVE PUPILS

PUPIL	READING AGE
A	135
B	147
C	118
D	156
E	145

STEP 1. *Transmuting test raw scores to subject-matter ages*

When a test has been given to a group of pupils, the raw scores resulting should be transmuted, according to the methods previously outlined, into subject ages such as reading ages, arithmetic ages, and the like. These ages should be expressed in months; if they have originally been determined in years and months, they should now be transmuted by means of Table XLI, p. 270.

If a reading test has been given to a group of pupils, let it be supposed that the reading ages are as given in Table LII.

[1] For a discussion of quotients and ratios, see pages 295–296.

STEP 2. *Determining the chronological ages*

The chronological ages of the pupils, expressed in months like the subject-matter ages, should be determined as outlined on page 95. The chronological ages for the five pupils of Table LII are as given in Table LIII.

TABLE LIII. CHRONOLOGICAL AGES OF PUPILS OF TABLE LII

PUPIL	CA
A	162
B	145
C	120
D	121
E	135

TABLE LIV. READING QUOTIENTS FOR PUPILS OF TABLES LII AND LIII

PUPIL	REA	CA	REQ
A	135	162	83
B	147	145	101
C	118	120	102
D	156	121	129
E	145	135	107

STEP 3. *Dividing the subject age by the chronological age to find the achievement quotient*

The achievement quotient is found by dividing the subject age of the pupil by his chronological age, according to the following formula:

$$(100 \times) \frac{\text{Subject age } (SA)}{\text{Chronological age } (CA)} = \text{achievement quotient } (AQ).$$

For the five pupils measured in reading, shown in Tables LII, LIII, the achievement quotients are as shown in Table LIV. In this table the achievement quotient for Pupil A was found by substituting in the foregoing formula as follows:

$$(100 \times) \frac{135\ (ReA)}{162\ (CA)} = 83\ (ReQ).$$

STEP 4. *Combining subject ages or average subject quotients to determine educational quotients*

This step is necessary only when a general-achievement quotient in several subjects in combination is desired for each pupil. When this is desired, a composite quotient can

be found by combining the subject ages (such as the ages for a single pupil in arithmetic, reading, and spelling) and dividing the composite by his chronological age. For example, if a pupil has achieved an arithmetic age of 140 months, a reading age of 155 months, and a spelling age of 135 months, these three ages would be averaged.

$$(100 \times) \frac{140 + 155 + 135}{3} = 143\ (EA).$$

This educational age would then be divided, according to the achievement-quotient formula, by the chronological age, 152, as follows:

$$(100 \times) \frac{143\ (EA)}{152\ (CA)} = 94\ (EQ).$$

Achievement quotients can be interpreted in the same way as described for intelligence quotients, according to the information contained in Table XLIV, p. 272. In general a quotient above 100 indicates a previous rate of learning greater than normal, a quotient of 100 a normal previous rate of learning, and a quotient of less than 100 a retarded previous rate of learning for the individual.

Determination of the B score in education. The B score in education is analogous to the achievement quotient and is a measure for previous rate of learning in the T scale. The B score may be found for any subject, as in the case of the achievement quotient, in which case it is called the B score in reading (B_r), the B score in arithmetic (B_a), and the like; or a combination of B scores may be found, which may be called the B score in education (B_e). Any of these B scores may be found as follows:

STEP 1. *Determining T scores in subject matter*

The determination of the T score in the particular type of subject matter to be used for the B score, or the several types to be used for the combined educational score, is the

first step in the determination. When a test has been given (for example, in reading), the raw scores are transmuted into T scores according to the method described on pages 240–246. The T_r scores for five pupils are shown in Table LV.

STEP 2. *Determining the chronological age*

The exact chronological ages of the pupils in terms of years and months should next be determined as outlined on page 95. When the ages are determined for the five pupils of Table LV they appear as in Table LVI.

TABLE LV. T SCORES FOR FIVE PUPILS IN READING

PUPIL	T SCORE
A	43
B	56
C	39
D	48
E	62

TABLE LVI. CHRONOLOGICAL AGES OF PUPILS OF TABLE LV

PUPIL	CA
A	13–4
B	12–2
C	10–6
D	14–3
E	13–7

STEP 3. *Determining the B correction for each pupil*

The B correction is found in terms of the chronological age for each pupil by means of Table XLVI, p. 275. Pupil A, who has a chronological age of 13–4, has, according to Table XLVI, a B correction of − 3. The B corrections for the remaining pupils are as shown in Table LVII.

TABLE LVII. B CORRECTIONS FOR PUPILS OF TABLE LVI

PUPIL	CA	BCR
A	13–4	−3
B	12–2	+1
C	10–6	+9
D	14–3	−7
E	13–7	−4

TABLE LVIII. B SCORES FOR FIVE PUPILS IN READING

PUPIL	T	BCR	Br
A	43	− 3	40
B	56	+ 1	57
C	39	+ 9	48
D	48	− 7	41
E	62	− 4	58

STEP 4. *Adding the B correction to the T score*

When the B correction for each pupil is added to his T score, the result is the B score. The formula and technic are described on page 276. For the pupils for whom B corrections were found in Table LVI the B scores in reading are as found in Table LVIII on page 283.

STEP 5. *Combining all the T scores for a single pupil, dividing by the number of scores, and applying the B-score formula to determine a general B score in education*

This step need be taken only when a combined B score is desired. The process has been outlined above and consists merely in finding an average, or a composite T score, which can then be treated as though it were a T score in a single subject. For illustration let it be supposed that a pupil has a T score in reading of 42, a T score in arithmetic of 49, and a T score in spelling of 37. Then

$$\frac{42 + 49 + 37}{3} = 43\ (T_e).$$

If the pupil is aged 12–8,the B correction, as found in Table XLVI on page 275, is − 1.

$$(T_e)\ 43 - (BCr)\ 1 = 42\ (B_e).$$

The interpretation of B scores, either the scores for a particular subject matter or the composite scores, may be made as with the B_i scores previously described, by means of Table XLIX or L. B scores in education can also be transmuted into educational quotients by means of Table LI. In general, a B_e score above 50 indicates a previous rate of learning greater than normal for individuals of the pupil's own age; a score of 50 a normal rate, and a score of less than 50 one less than normal.

Measures of probable future rate of learning. The probable future rate of a pupil's learning is determined by how closely

he can be helped to approach his potential rate of learning. This in turn is partly conditioned by what he actually has done in the past. The measures for the determination of probable future rate of learning are therefore combinations of measures which indicate potential and previous rates.

In the age scale the measure of probable future rate of learning is known as the promotion quotient (PrQ); in the T scale it is known as the B score for promotion (B_p).

Determination of the promotion quotient. The promotion quotient is a relationship between the chronological age and a composite of the general educational age (EA) and the mental age (MA); therefore it yields a quotient similar in characteristics to the quotient previously considered and shows what can probably be expected from a pupil as a probable future rate of learning.

STEP 1. *Determining the mental age*

The mental age is determined by transmuting the raw scores from an intelligence test into mental ages in months, as shown on page 260.

STEP 2. *Determining a general educational age*

A general educational age is determined from a series of educational tests by combining the subject ages for a pupil into a single composite age, expressed in months, as described on page 281. When this step has been completed, the mental and educational ages for five pupils are as listed in Table LIX. The first column shows the names of the pupils, the second their mental ages, the third their reading ages (ReA), the fourth their arithmetic ages (ArA), the fifth their spelling (SpA), and the sixth the composite ages in general educational achievement that were obtained by applying the formula

$$\frac{ReA + ArA + SpA}{3} = EA.$$

TABLE LIX. EDUCATIONAL AGES FOR FIVE CHILDREN

PUPIL	MA	REA	ARA	SPA	EA
A	144	151	136	150	146
B	167	163	180	142	162
C	142	120	125	121	122
D	100	97	85	90	91
E	135	132	140	134	135

STEP 3. *Obtaining a promotion age*

A promotion age is a combination of all available ages which are significant in determining a pupil's status. In this case the ages available are the mental and educational ages. In Chapter XIII, p. 355, a technic will be shown for determining another significant age which might be added to those mentioned above. Here the promotion age is found from the formula

$$\frac{MA + EA}{2} = PrA.$$

The promotion ages for the five pupils of Table LIX are as shown in Table LX, where they are derived from the data there given.

TABLE LX. PROMOTION AGES FOR FIVE PUPILS

PUPIL	MA	EA	PRA
A	144	146	145
B	167	162	165
C	142	122	132
D	100	91	96
E	135	135	135

TABLE LXI. SHOWING PROMOTION QUOTIENTS FOR FIVE PUPILS

PUPIL	CA	PRA	PRQ
A	162	145	90
B	145	165	114
C	120	132	110
D	121	96	79
E	135	135	100

STEP 4. *Determining the chronological age*

The chronological ages of the pupils are determined as shown on page 95. For the five pupils under consideration

the chronological ages are as shown in Table LXI on page 286 in the column labeled "CA."

STEP 5. *Determining the promotion quotient*

The promotion quotient, like the other quotients previously discussed, is a relationship between the subject age and the chronological age. Here it is a relation existing between the promotion age and the chronological age, and is found through the formula

$$(100 \times) \frac{PrA}{CA} = PrQ.$$

In the case of Pupil A of Table LXI, whose promotion age is 145 and whose chronological age is 162, the substitution becomes

$$(100 \times) \frac{145\ (PrA)}{162\ (CA)} = 90\ (PrQ).$$

When the substitutions have been made for the remaining pupils, the results are as found in Table LXI, which contains all the essential data.

The interpretation of the promotion quotient is similar to that of other quotients and shows what may be expected of a pupil in terms of probable future rate. When the quotient is 100 or thereabouts, the probable future rate of learning is normal; when it is below 100, the rate of the pupil will probably be less than that of pupils of his own age; when it is greater than 100, the rate will probably be greater than for pupils of his own age. As will be shown later, the quotient is therefore of use in helping to determine the type of group in which a pupil would be best placed.

Determination of the B score for promotion. Like the promotion quotient, the B score for promotion in a measure which combines the measures of potential and previous rates of learning into one inclusive measure which shows the probable future rate of learning.

STEP 1. *Determining the B score in intelligence* (B_i)

The B score in intelligence is found as described on page 272, by giving an intelligence test, determining the T_i score, and from that obtaining the B_i score.

STEP 2. *Determining the B score in education* (B_e)

The B score in education is found as described on page 282, by transmuting the T scores on educational tests into B scores, and from the B scores obtaining a composite B score.

Table LXII shows the derivation of B_i and B_e scores for five pupils. The first column gives the names of the pupil, the second their chronological ages, the third their T scores in intelligence (T_i), the fourth their T scores in reading (T_r), the fifth their T scores in arithmetic (T_a), the sixth their T scores in spelling (T_s), the seventh the combined T_e scores, the eighth their B corrections (BCr), and the ninth and tenth the B scores for the T_i and T_e scores in the third and seventh columns (B_i and B_e).

TABLE LXII. DERIVATION OF B_i AND B_e SCORES FOR FIVE PUPILS

PUPIL	CA	T_i	T_r	T_a	T_s	T_e	BCr	B_i	B_e
A	10–6	40	44	47	43	45	+9	49	54
B	14–2	52	39	42	37	39	−6	46	33
C	10–7	46	48	47	50	48	+8	54	56
D	8–9	37	38	34	30	34	+20	57	54
E	12–3	48	38	40	43	40	+1	49	41

STEP 3. *Determining the B score for promotion* (B_p)

The B score for promotion is found through the formula

$$\frac{B_i + B_e}{2} = B_p.$$

For Pupil A the substitution becomes

$$\frac{49\ (B_i) + 54\ (B_e)}{2} = 52\ (B_p).$$

When the substitutions have been made for the remaining pupils of Table LXII, the results are as shown in Table LXIII.

The B_p may be interpreted as indicating that in pupils with B_p ratings above 50 the rate of learning is greater than is normal for pupils of the same chronological age, while ratings below 50 are less than is normal for pupils of the same chronological age. When taken in conjunction with the effort score, which is discussed in Chapter XII, the B_p is a useful rating for aiding in the proper classification or promotion of pupils.

TABLE LXIII. DETERMINATION OF B_p FOR FIVE PUPILS

PUPIL	B_i	B_e	B_p
A	49	54	52
B	46	33	40
C	54	56	55
D	57	54	56
E	49	41	45

Chapter summary. Three measures of the rate of learning are of significance in the knowledge of any pupil. A pupil's potential rate of learning is the rate of learning which is possible for him in terms of his inherited ability. In the age scale the measure is the intelligence quotient, known as the IQ. The IQ is found from an intelligence test by dividing the mental age by the chronological age. In the T scale the measure of potential rate of learning is the B score in intelligence, known as B_i. This measure is found by transposing the T score on an intelligence test for any pupil into a B score through the addition or subtraction of a B correction for the chronological age of the pupil. The B score then becomes a measure in terms of T units, where the score of 50 represents the normal rate for pupils of any given chronological age.

The previous rate of learning is that rate which a pupil has actually achieved in learning as a result of his ability, his habits, and teaching. In the age scale the measures are the achievement quotient and the educational quotient, known

as the AQ and the EQ. The AQ is found by dividing the subject-matter age of a pupil by his chronological age; the EQ is found by making a composite of several subject-matter ages and dividing that by the chronological age. In the T scale the measures are the B score in subject matter and the B score in education. The B score in some phase of subject matter is found by adding (regarding signs) the B correction for the age of the pupil to the T score in that phase of subject matter. The B score in education (B_e) is found by making a composite of several T scores in different phases of subject matter and adding to that composite (regarding signs) the B correction for the pupil's chronological age.

The probable future rate of learning is a combination measure which includes both the potential and the previous rate of learning. In the age scale the measure is the promotion quotient, known as PrQ. This measure is found by making a composite of the mental and educational ages, which is then divided by the pupil's chronological age. In the T scale the measure is the B score for promotion, known as the B_p. It is determined by finding a composite of the B score in intelligence and the B score in education.

In general it may be said that the quotients are always some relation to the pupil's chronological age. When they are above 100, they indicate a rate greater than that for normal pupils of the same age; when they are below 100, they indicate rates less than normal for pupils of the same age. The B score for a pupil is always in T units, reduced to a scale where the mean is that of pupils of the same chronological age. Therefore when a B score is above 50, it indicates a rate of learning greater than that of normal pupils of the same age; when it is below 50, it indicates a rate less than that of normal pupils of the same age.

SELECTED BIBLIOGRAPHY

FENTON, N., and WORCESTER, D. A. An Introduction to Educational Measurements, chap. ii. Ginn and Company, Boston, 1928.

GREGORY, C. A. Fundamentals of Educational Measurement, chaps. iii–iv. D. Appleton and Company, New York, 1922.

MCCALL, W. A. How to Measure in Education, pp. 79–86. The Macmillan Company, New York, 1922.

PINTNER, R. Intelligence Testing, chap. v. Henry Holt and Company, New York, 1923.

TERMAN, L. M. The Intelligence of School Children. Houghton Mifflin Company, Boston, 1919.

TERMAN, L. M. The Measurement of Intelligence. Houghton Mifflin Company, Boston, 1916.

THORNDIKE, E. L. The Measurement of Intelligence. Teachers College Bureau of Publications, New York, 1927.

TRABUE, M. R. Measuring Results in Education, chap. xviii. American Book Company, New York, 1924.

CHAPTER XII

MEASURES OF EFFORT

Relation of effort to achievement. In the learning process there are three elements which are of importance to the teacher. There is, first, that inborn possibility to learn, variable as between individuals but present in each. There is, secondly, the achievement of an individual. This is the actual learning which was discussed under the term "achievement" and which was also found to be subject to variations in terms of rate as between individuals. A third element, difficult to name, also variable as between individuals, may be thought of as the will to learn or the will to do.

No matter what the inherent ability may be, if the will to do is not present the talent may be wasted. No matter how great the achievement of an individual may be, unless it is comparable with the ability it is not success. This will to do is what is meant by effort, and the measurement of that effort is one of the great additions to modern measurement.

It has been seen that achievement measures are those which measure the actual extent of achievement. It has also been seen that ability measures are those which measure the inherent ability. Effort measures are those which measure the extent to which achievement is made in terms of that ability.

Importance of effort in the classroom. Without question the traditional desire of those in charge of schools and school work was to have the school work result in definite accomplishment. As stated this is trite and not far different from what a teacher would accept today as the objective of school

work. But the emphasis that was laid on the idea traditionally was an entirely different emphasis than that which is in vogue today. The accomplishment that was desired was almost entirely in terms of subject matter. As a result subject matter was divided into units, arbitrarily and empirically, and all pupils were expected to reach the same degree of accomplishment. School methods were based upon this idea, school organizations were built up upon it, and schoolbooks were founded upon it. The net result was the recognition of the pupil as a member of a group but not as an entity in any way variable from that group.

If extremes may be cited, two difficulties have come from this idea. On the one hand, there is a definite retardation of the more capable; on the other, there is a definite attempted acceleration of the less gifted.

With the supernormal pupil whose work is usually entirely acceptable in terms of the work of the group as a whole of which he is a part, there is an implied satisfaction in a degree of achievement far less than the ability would warrant. With the subnormal child, where the achievement is undoubtedly inferior to that of the group as a whole of which he is a part, there is an implied dissatisfaction with the degree of work which is in reality greater than one would have the right to expect. This tendency can be illustrated in Figure 79.

If the vertical lines *A* and *B* represent the relative abilities of two pupils, the line *X* representing the extent of Pupil A's achievement, and the line *Y* representing the extent of the achievement of Pupil B, it can be easily seen that the total achievement of Pupil A is greater than that of Pupil B. If the horizontal line represents a point of what is usually thought of as satisfactory achievement for the grade in which these pupils are, it can also be seen that Pupil A has exceeded that point, whereas Pupil B has not reached it. Under these circumstances Pupil A is praised and given a good grade for

satisfactory work, and Pupil B is blamed for his lack of achievement and is given a poor grade.

Pupil A is told that his work is satisfactory, and by high grades is given to understand that he need not work harder; Pupil B is told that his work is unsatisfactory, and is practically forced to work harder.

The real facts in the case, if the pupils be considered in the light of modern knowledge of individuals, are that the work

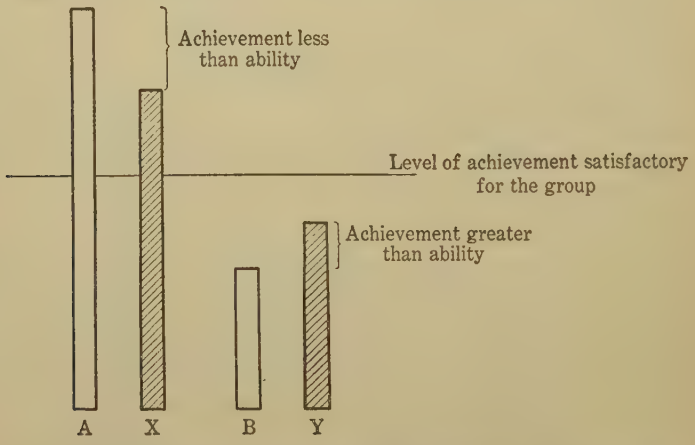


FIG. 79. Comparison of ability and achievement of two pupils

of Pupil A, regardless of the extent of his absolute achievement, is unsatisfactory. He has done less than his ability would warrant. The work of Pupil B, regardless of the extent of his absolute achievement, is more than satisfactory. Pupil A has done less than he might have done, and Pupil B has done more. Pupil B should receive praise for his achievement, considering his ability, and Pupil A should be helped to achieve the promise of his ability.

Essentials in the measurement of effort. One necessity in the measurement of effort, as has been implied, is the meas-

ure of achievement. This measure will show what an individual actually has accomplished. A second necessity is some measure of ability for the same individual. This will show what the individual might have accomplished.

In addition it is necessary that these two measures should be expressed in the same kind of units. The measure of effort is one which shows the relationship of achievement to ability, and in order to do this the two measures must be comparable. In addition to being comparable in terms of units the two measures must also be comparable in terms of time. They must both represent the status of the pupil at the same time. An ability as expressed for a given time cannot be compared with an achievement of six months later: it must be compared with an achievement as measured at the same time as the ability is measured, or it must be compared with an achievement where the measure of ability has been corrected for the time at which the achievement was measured.

These four elements — a measure of ability, a measure of achievement, with the two measures agreeing both in the form of unit and in time — are the essentials of the measurement of the individual's degree of effort.

The effort measure in the age scale. In the age scale the effort measure is the achievement ratio. This measure has been variously called the achievement quotient, the accomplishment quotient, and the accomplishment ratio. It has been noted that a measure of rate, not effort, is called the achievement quotient. This is consistent with present practice, although when the effort measure was first proposed[1] it was termed a quotient.

An educational *quotient* is the relationship between the achievement of a pupil and that of other pupils of the same

[1] R. Franzen, *The Accomplishment Quotient* Teachers College Bureau of Publications, Teachers College, New York City, 1921. William A. McCall, *How to Measure in Education* (The Macmillan Company, New York, 1923), pp. 85–87.

chronological age. In the formula for determining the quotient, therefore, the denominator of the fraction representing age and achievement is always the chronological age.

An educational *ratio* is a relationship between the achievement of a pupil and his own ability. In the formula for determining the ratio the denominator of the fraction representing the relationship between ability and achievement is always the mental age.

In the age scale the achievement ratio is determined as follows:

STEP 1. *Determining the mental age*

The mental age of each pupil should be determined for the time that the test of achievement was given. This mental age can be derived from an intelligence test, as has been previously described, or, if an intelligence test has already been given to the pupils, it can be derived from the scores or mental ages which resulted from the test.

TABLE LXIV. MENTAL AGES OF FIVE PUPILS

PUPIL	MENTAL AGE	MENTAL AGE IN MONTHS
A	12–6	150
B	11–8	140
C	12–9	153
D	9–6	114
E	10–2	122

In the first case it is necessary to give some form of group intelligence test, to transmute the raw scores by means of a table usually furnished with the manual of directions into an age score for each pupil, and to list the final age scores.

If an intelligence test has already been given, it is only necessary to determine the intelligence quotients for the pupils and to multiply the chronological age of each pupil at the time of giving the achievement test by his quotient. This will result in a series of mental ages corrected for the time of the giving of the achievement test. Such a group of mental ages for five pupils is shown in Table LXIV.

STEP 2. *Determining the educational achievement age*

The next step is to determine the achievement age of each pupil for whom an effort score is to be derived. To do this a test should be given in the subject for which an effort score is desired, and the raw scores on the test should be transmuted into age scores by the method described on page 212.

Such a group of achievement scores for the same group of pupils included in Table LXIV is shown in Table LXV. In each case the ages must be turned into months by means of Table XLI, so that they can be handled more easily.

TABLE LXV. READING ACHIEVEMENT AGES OF SAME FIVE PUPILS AS SHOWN IN TABLE LXIV

PUPIL	READING AGE	READING AGE IN MONTHS
A	12–0	144
B	10–7	127
C	12–9	153
D	10–7	127
E	10–3	123

STEP 3. *Applying the achievement-ratio formula*

The third step is to apply the achievement-ratio formula in the case of each pupil. This will give an effort measure for each pupil. The formula for the achievement ratio is

$$(100 \times) \frac{\text{Achievement age}}{\text{Mental age}} = \text{Achievement ratio.}$$

Pupil A, Table LXIV, has a reading age of 144 months. His mental age, corrected to the date of the reading test, is 150 months. Applying the formula to these data gives

$$(100 \times) \frac{144\ (RA)}{150\ (MA)} = 96\ (AR \text{ for Pupil A}).$$

It will be noted that the ratio is multiplied by 100 in the same way as quotients were multiplied, in order to remove the decimals. When the achievement ratios for the other pupils have been found according to the formula given above, they are as shown in Table LXVI.

Interpretation of achievement ratios. Achievement ratios give a rating of a pupil which is somewhat more significant in individual cases than is the achievement quotient previously described. The ratio shows what a pupil has done in terms of his own ability, whereas the quotient shows what he has done in comparison with normal pupils of his own age. Thus, if an achievement ratio is above 100, it indicates that the pupil has done better than one would expect for a pupil of his ability; if it is approximately 100, it indicates that he has equaled the promise of his ability; if it is below 100, it indicates that he has done more poorly than his ability would warrant.

TABLE LXVI. ACHIEVEMENT RATIOS FOR FIVE PUPILS IN READING

PUPIL	MA	REA	AR
A	150	144	96
B	140	127	91
C	158	158	100
D	114	127	111
E	122	123	101

The achievement ratio is a measure which should be used by a teacher when working with results from standard tests in the age scale, before making an absolute decision as to the quality of the work of any pupil. The pupil who is apparently doing well and who is at the top of the class may be found to have a low achievement ratio, and should therefore be encouraged to do better; the pupil apparently at the bottom of the class may be found to be exceeding, even with his apparently poor work, the standard set for him by his ability and should either be restrained or reclassified. A discussion of these problems will be found in Chapters XIII–XIV.

The effort measure in the T scale. In the T scale the measure of effort is called the F score. This is a somewhat easier measure to find than the achievement ratio, since it can be found by simple addition and subtraction rather than by division. In the F score the attempt is made to reduce the effort of a pupil to a scale similar to the T scale and, at the same

time, to show the relationship of that effort to what would be normal for one of his ability.

If a pupil has a B_i of 57, it means that his potential rate of learning is 7 T units greater than that of normal pupils of his own age. This then becomes the measure of the rate at which the pupil should learn in comparison with normal pupils. If his actual rate of learning in reading (B_r) is 43, superficially it would appear that he was actually learning at

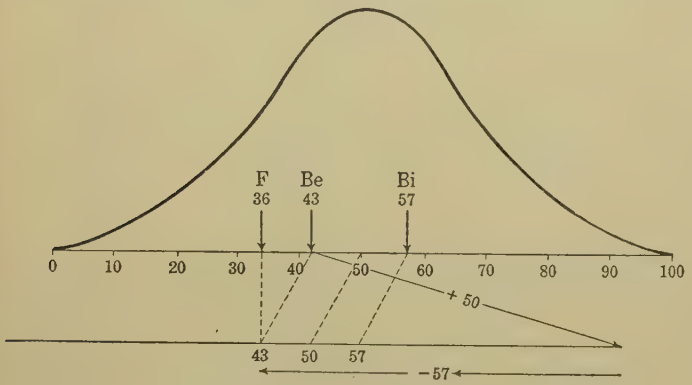


FIG. 80. Derivation of the F score

a rate 7 T units less than normal, which is true for normal pupils of his age but is not true for him, since his ability is 7 T units above normal. It is necessary, then, in determining effort to take into account not the achievement of the pupil in relation to normal pupils of his own age, but rather the achievement to be expected from one of his ability or potential rate of learning. The process of finding the effort score is one which reduces either the T_i or the B_i of a pupil to 50 (which in all other scales has been average or normal), and then relocates the T or B score in subject matter or education on that basis. Fig. 80 shows graphically how this is done for B scores. As a first step 50 is added to the B_e score, and from

this sum the B_i (57) is subtracted. This has the effect of reducing the B_i to 50 and shows at what point the B score in education lies on this new scale.

The formula for finding the effort score is as follows:

$$(T_e + 50) - T_i = \text{F score.}$$
$$(B_e + 50) - B_i = \text{F score.}$$

This formula applied to the case given above in terms of B scores shows the substitution as follows:

$$(43(B_r) + 50) - 57(B_i) = 36 \text{ (F score).}$$

This F score shows that this pupil who is apparently only 7 T units below the normal rate is actually 14 T units below the rate which is warranted by his own ability. The process for finding F scores is therefore as follows:

STEP 1. *Determining the T_e or B_e for each pupil*

The B score in intelligence is found by giving an intelligence test, by turning the raw scores on the test into T scores, by finding a B correction for each pupil, and by adding (regarding signs) this B correction to the T score for each pupil, finding a B_i score.[1]

STEP 2. *Determining the T_e or B_e for each pupil*

The B score in education may be determined for a pupil in specific forms of subject matter, such as reading, arithmetic, spelling, and the like, or it may be determined for a composite of these subjects. The process for finding the F score in reading or arithmetic and the like is the same as that given below for finding the F score in education, save that where T_e and B_e are used, T_r, T_a, T_s or B_r, B_a, B_s, and the like should be substituted. These B scores are derived from the T scores by adding (regarding signs) the B corrections shown on page 275.

[1] See pages 272–277.

In Table LXVII are shown the B_e and B_i scores of the five pupils as derived from Table LXII (p. 288) and found by the method here referred to.

STEP 3. *Applying the F-score formula*

The formulas for finding F score have been given above as

$$(T_e + 50) - T_i = F.$$
$$(B_e + 50) - B_i = F.$$

In the case of Pupil A, who has a B_e of 54 and a B_i of 49, the formula becomes

$$(54\,(B_e) + 50) - 49(B_i) = 55.$$

When the F scores have been found in this way for the remaining pupils in Table LXVII, they stand as found in the last column of that table.

Interpretation of F scores. The F score, like the achievement ratio, shows the extent to which a pupil has achieved the promise of his ability. In the F scale, as in the other scales based on the T scale, 50 is normal. In this case normal effort means such an effort as is commensurate with the ability of the individual and therefore must be contrasted with the B score, which shows the relationship of the rate of growth of the pupil not in terms of himself, but in terms of the rate of learning of normal pupils of his own age.

TABLE LXVII. DERIVATION OF F SCORES FOR FIVE PUPILS

PUPIL	B_e	B_i	F
A	54	49	55
B	33	46	57
C	56	54	52
D	54	57	47
E	41	41	50

An F score of more than 50 indicates an effort on the part of a pupil which is greater than one should expect from one of his ability, an F score of approximately 50 indicates that a pupil has expended an effort to learn equivalent to his ability, and an F score of less than 50 indicates that the effort which

a pupil has expended is less than would be warranted by one of his ability. Thus the direction of correction is indicated.

In subsequent chapters the use of the F score in helping to determine the judgment of the teacher with respect to the pupil will be further expanded, and it will be shown that, regardless of the absolute placement of a pupil within a class, judgment as to his real success should be measured in terms of his achievement in terms of his ability, as well as in terms of his total achievement. In many cases a pupil may have a high achievement but a low effort score, showing that he should have achieved even more, whereas in other cases a pupil may have a relatively low achievement and still be doing all that could reasonably be expected of him.

Chapter summary. In addition to the inborn capacity to learn and the actual achievement in learning which a pupil possesses, there is a third factor of importance in teaching, and particularly of importance in judging the work of a pupil, which consists in the effort he makes, measured in terms of his ability to learn. The tendency in classroom procedure, where emphasis is placed upon the achievement and progress of the group, is to set standards of achievement for the group. As a result, unless some measure of the effort expended by pupils is made, no true judgment of the actual quality of achievement can be made.

The effort measure is a relationship between the ability of the pupil and his actual achievement. In the age scale it is called the achievement ratio; in the T scale, the F score.

The achievement ratio is found by reducing a mental age to months and by reducing some subject age or a composite of subject ages also to months. The direct ratio between these two ages, using the mental age as the denominator and the educational age as the numerator, is the achievement ratio. In practice the achievement ratio is multiplied by 100 in order to remove the decimal point. When this is done a ratio above

100 indicates effort greater than is expected, a ratio of approximately 100 indicates work normal for a pupil of that ability, and a ratio of less than 100 indicates less effort than should be expected from this pupil.

The F score is found through reducing the T or B score in intelligence to 50 and showing the relationship, in T units, to this revised T_i or B_i score. It is found by adding 50 to the T or B score in subject matter or to a composite T or B score in education, and subtracting from that the T or B score in intelligence. When this has been done, an F score above 50 indicates an expenditure of effort greater than could be expected from a pupil of the ability given, an F score of approximately 50 indicates a degree of effort which is normal for this pupil, and an F score of less than 50 indicates a degree of effort which is less than is warranted by his ability. Later chapters consider the classroom utility of the effort scores.

SELECTED BIBLIOGRAPHY

FENTON, N., and WORCESTER, D. A. An Introduction to Educational Measurements, pp. 55–57. Ginn and Company, Boston, 1828.

FRANZEN, R. The Accomplishment Ratio (Teachers College Contributions to Education No. 125). Teachers College Bureau of Publications, New York, 1922.

MCCALL, W. A. How to Measure in Education, pp. 85–87. The Macmillan Company, New York, 1922.

MONROE, W. S. The Illinois Examination, *University of Illinois Bureau of Educational Research Bulletin No. 6* (October 31, 1921), Vol. XIX, No. 91.

ODELL, C. W. Traditional Examinations and New-Type Tests, p. 124. The Century Co., New York, 1928.

OTIS, A. S. Statistical Method in Educational Measurement, pp. 4, 172, 337. World Book Company, Yonkers-on-Hudson, 1925.

RUSSELL, C. Classroom Tests, chap. xv. Ginn and Company, Boston, 1926.

TRABUE, M. R. Measuring Results in Education, p. 445. American Book Company, New York, 1924.

WILSON, G. M., and HOKE, K. J. How to Measure (Revised Edition), pp. 369–372, 537. The Macmillan Company, New York, 1928.

PART IV. USES

CHAPTER XIII

CLASSIFICATION OF PUPILS WITHIN A SCHOOL

The problem of classification. Classification presents one of the more difficult problems connected with dealing with pupils in schools. Economic and social requirements necessitate the grouping of pupils in classes and grades. The pupils in these classes and grades, however, form a variable group, and differ in many respects from one another. They differ in age, in background, in ability, in response to teaching, in rate of progress, and in level of attainment. With such variations in the character of the individual pupils in a group, and with the added knowledge that the fact of variability will exist, the problem of classifying such pupils in proper places where they can profit most from instruction is both a social need for the pupil and an economic need for the school system.

In all probability the greatest light that has been thrown on the problems of classification has come through the medium of testing, and standard tests used with discrimination and understanding have formed the basic solution.

Classification within a school must take into account these various factors of differences as among pupils and, so far as possible, place each pupil in such grade and with such other pupils as will give him the greatest possibility for development while they are taught in classes. Such a need necessitates the determination of as many points of knowledge of

a pupil as it is possible to find, and the reduction of all these elements to some single grade unit which will show in what school unit he is best placed. The actual achievement of the pupils in many phases of their work is one important element. A second element is the ability of the pupils; this is an indication of the rate at which they ought to progress. A further element that should be considered is the degree to which they have achieved in terms of their ability. The classification of a school involves all these elements, plus others, which reënforce the judgments thus obtained. In general, classification means the placing together of pupils of equal educational status who progress at approximately equal rates.

Process of classification. To classify pupils properly within a school system, it is necessary to find some ratings which will place a pupil in the grade or class where he will be able to do his best work. A pupil's best work is based partly on his ability at the moment, the status of his achievement, his physical, emotional, and (possibly) social status, and his willingness to do.

The determination of a pupil's proper classification, therefore, means finding — so far as possible — ratings for each of these elements, giving each of them a weighting commensurate with its worth for classification or promotion, turning these ratings into measures which are comparable with one another, and combining them or using them in determining the pupil's grade status for promotion. The standard measures — achievement, ability, rate, and effort measures which have been described — make possible this use of the results of testing and can therefore be used for classification data. The T scale, the age scale, and the age and grade scales combined may be used for classification, as is described later in this chapter. It is merely a means for combining the entire knowledge of a teacher in a useful way.

Classification of a school by means of the T scale. When a school is to be classified by means of the T scale, the steps in the classification are as follows:

STEP 1. *Selecting an ability test*

When an entire school is to be reclassified in terms of standard tests, the first step is to determine the ability of each pupil in the school. This can best be done by means of a group intelligence test, like that described in Chapter X. Any good language group intelligence test can be used and should be given to all the pupils in the school who are capable of taking it. In most cases it will be found possible to give such a test to all the pupils from the third grade through the eighth. Third-grade pupils as a rule are measured successfully in ability by such language tests as the Haggerty, Otis, National Intelligence, and like tests.

Below the third grade, where there are real language difficulties, it is wise to give some other form of test (preferably a non-language test) suited to such pupils. In the classification to be described in illustration of the technic of classification the National Intelligence Test was given to all pupils in the third grade and above, and the Pintner-Cunningham Primary Mental Test was given to all those in the first and second grades.

When the results of these tests had been found through scoring and re-scoring, the raw scores on the tests were changed into T scores for the purpose of making them comparable with other scores in other tests which were contemplated. The first and fifth grades are used to illustrate the process throughout, and the following results were obtained from the use of the ability tests in these grades. Table LXVIII, second column, shows the raw scores on the tests for the first grade for each of the pupils named in the first column; the third column shows the T scores in intel-

TABLE LXVIII. INTELLIGENCE SCORES FOR FIRST GRADE

(Pintner-Cunningham)

PUPIL	RAW SCORE	T_i	PUPIL	RAW SCORE	T_i
A	34	16	P	29	11
B	41	24	Q	32	13
C	33	15	R	28	8
D	28	8	S	22	−2
E	17	−5	T	22	−2
F	47	29	U	36	19
G	37	20	V	23	0
H	11	−12	W	29	9
I	22	−2	X	35	17
J	40	23	Y	30	11
K	38	20	Z	29	9
L	40	23	AA	39	22
M	27	6	BB	25	2
N	34	16	CC	46	28
O	37	20	DD	51	32

TABLE LXIX. INTELLIGENCE SCORES FOR FIFTH GRADE (N.I.T.)

PUPIL	RAW SCORE	T_i	PUPIL	RAW SCORE	T_i
A	72.0	39	T	93.0	45
B	106.3	49	U	124.9	55
C	119.3	53	V	90.3	44
D	100.3	47	W	93.4	45
E	66.9	38	X	133.8	57
F	115.2	52	Y	114.1	52
G	145.5	62	Z	141.7	60
H	124.4	55	AA	79.6	41
I	78.1	41	BB	95.1	46
J	109.8	50	CC	154.4	66
K	127.8	55	DD	100.8	47
L	120.5	53	EE	96.4	46
M	128.7	56	FF	87.6	43
N	77.8	41	GG	75.0	40
O	114.0	52	HH	130.0	56
P	82.8	42	II	79.4	41
Q	107.4	49	JJ	84.7	43
R	117.8	52	KK	91.8	44
S	147.6	63			

ligence (T_i) obtained as a result of transmutation of the scores by the method described in Chapter X.

Table LXIX on page 307 shows in the first column the name of each pupil in a fifth grade, and in the second column the raw score made on the National Intelligence Test; by transmutation according to the method described on page 260, in the third column are the T scores in intelligence (T_i) corresponding to the raw scores given. The transmutation table used in this case is given in Table XXXVII and is a T table made up especially by Dr. McCall for use with the National Intelligence Test.

When these scores have been obtained (and, as explained later, this can be done simultaneously with the development of other scores), they should be carefully preserved for inclusion in the classification tables which are to be constructed.

STEP 2. *Selecting and giving achievement tests*

The second step in the process of reclassification is the selection and giving of a battery of achievement tests. These tests should be selected with care, so as to achieve the desired results and, in so far as possible, to cover the entire range of grades to be reclassified. The reclassification increases in value when the same factors are used in judging all the pupils, regardless of their grades. It is possible to give a single test covering a wide range of subject matter, such as the Stanford Achievement or the Illinois Examination, which were described in Chapter IV, or it is possible to select from among the great variety of achievement tests those which may be considered as being indicative of the general educational progress of the pupils. In the situation that is being used here for illustration it was desirable to give as few tests as possible in order to minimize the labor, to have the tests which were selected offer a range of measurement of both skills and knowledge, and also to find variety in the testing

and the testing methods, since the testing was under the observation of students. Other combinations of tests may be found more suitable in other situations, and when reclassification is contemplated, not only the primary use of the tests for reclassification should be considered, but also their secondary uses for diagnosis or supervision.

The tests selected for this reclassification were, for the third grade and up, the Thorndike-McCall Reading Scale, Form 2, the Woody-McCall Arithmetic Test, Mixed Fundamentals, Form 1, and the Morrison-McCall Spelling Test, List 4. The reading scale was selected because it was felt that through reading ability could be determined the general ability of the pupils in those subjects which did not range throughout the grades, where considerable importance must be placed upon reading, as in geography, history, and the like. The arithmetic test was selected as a measure of the pupils' ability in exact school skills, and the spelling test was utilized somewhat as a check upon the reading, because of the presence in the school of a fairly large proportion of pupils of foreign extraction. Many other tests of great value might have been given, such as tests in written composition, in handwriting, and the like, but it was felt that those selected offered as wide a range as was needed and also offered splendid norms for the transmutation of scores.

In the first and second grades it was decided to use one test in reading, rather than to attempt to carry the arithmetic and spelling to those grades. The test selected was the Haggerty, Sigma 1, a sample of which was shown on page 108.

A time was set for giving each of these tests throughout the entire school. The reading test was given in all the grades at the same hour, and a different hour was selected for giving the arithmetic and spelling tests.

When all the tests had been given, they were scored and checked, and the raw scores in each case were transmuted

TABLE LXX. READING SCORES FOR FIRST GRADE

(Haggerty, Sigma 1)

PUPIL	RAW SCORE	T_r	PUPIL	RAW SCORE	T_r
A	6	12	P	17	23
B	17	23	Q	11	17
C	2	8	R	2	8
D	13	19	S	2	8
E	2	8	T	4	10
F	21	27	U	16	22
G	7	13	V	4	10
H	3	9	W	17	23
I	5	11	X	7	13
J	11	17	Y	19	25
K	5	11	Z	5	11
L	26	31	AA	8	14
M	4	10	BB	13	19
N	2	8	CC	28	33
O	21	27	DD	31	36

TABLE LXXI. READING SCORES FOR FIFTH GRADE

(Thorndike-McCall)

PUPIL	RAW SCORE	T_r	PUPIL	RAW SCORE	T_r
A	17	42	T	14	38
B	20	47	U	25	58
C	26	60	V	12	36
D	18	43	W	20	47
E	16	40	X	25	58
F	18	43	Y	23	53
G	25	58	Z	23	53
H	23	53	AA	12	36
I	20	47	BB	16	40
J	23	53	CC	23	53
K	21	49	DD	22	51
L	23	53	EE	22	51
M	22	51	FF	16	40
N	14	38	GG	17	42
O	16	40	HH	22	51
P	16	40	II	16	40
Q	18	43	JJ	18	43
R	23	53	KK	21	49
S	28	67			

into T scores by the methods described on page 244. The raw scores in reading when transmuted became T scores in reading (T_r), the raw scores in arithmetic when transmuted became T scores in arithmetic (T_a), and the raw scores in spelling when transmuted became T scores in spelling (T_s).

After transmutation the record sheets in reading, arithmetic, and spelling for the first and fifth grades, for which the intelligence test results were given previously, are as given in Tables LXX–LXXIII. Table LXX shows the raw scores and their T-score equivalents for the pupils in the first grade in the reading test.

Table LXXI shows the raw scores and their T-score equivalents for the pupils in the fifth grade in the reading test.

Table LXXII shows the raw scores and their T-score equivalents for the pupils in the fifth grade in the arithmetic test.

TABLE LXXII. ARITHMETIC SCORES FOR FIFTH GRADE

(Woody-McCall)

PUPIL	RAW SCORE	T_a	PUPIL	RAW SCORE	T_a
A	16	34	T	23	44
B	21	41	U	26	50
C	26	50	V	21	41
D	24	46	W	23	44
E	24	46	X	28	54
F	21	41	Y	28	54
G	24	46	Z	21	41
H	33	70	AA	21	41
I	25	48	BB	21	41
J	25	48	CC	22	42
K	22	42	DD	29	56
L	26	50	EE	14	32
M	21	41	FF	27	51
N	16	34	GG	26	50
O	21	41	HH	27	51
P	22	42	II	21	41
Q	25	48	JJ	19	38
R	30	59	KK	24	46
S	24	46			

Table LXXIII shows the raw scores and their T-score equivalents for the pupils in the fifth grade in the spelling test.

TABLE LXXIII. SPELLING SCORES FOR FIFTH GRADE

(Morrison-McCall)

PUPIL	RAW SCORE	T_s	PUPIL	RAW SCORE	T_s
A	25	43	T	27	44
B	33	49	U	43	60
C	41	57	V	31	48
D	29	46	W	29	46
E	22	40	X	33	49
F	18	37	Y	29	46
G	37	53	Z	31	48
H	35	51	AA	37	53
I	22	40	BB	26	44
J	27	44	CC	31	48
K	30	47	DD	35	51
L	36	52	EE	28	45
M	36	52	FF	28	45
N	29	46	GG	35	51
O	22	40	HH	30	47
P	23	41	II	30	47
Q	33	49	JJ	25	43
R	32	48	KK	32	48
S	47	67			

STEP 3. *Constructing a reclassification blank*

When the data concerning intelligence and achievement have been found, as indicated in the previous sections, it becomes necessary to formulate and to prepare a reclassification blank upon which the data can be placed and from which further information can be found. It is desirable to make the blank very carefully and to make each item on it clearly so that it can be easily used. As will be shown later, the blank can be made more elaborate than the one given here, and should be extended when the data are to be placed in the hands of a teacher for diagnostic or supervisory purposes. It is not necessary to explain this blank in detail at

this point, since each item will be explained as this discussion continues on the pages immediately following.

A sheet of paper should be prepared in the following manner: Horizontally there should be a sufficient number of ruled lines to allow one for each pupil in the class, as well as two or more lines for labeling at the top of the page. Vertical lines should then be drawn across these horizontal lines, so as to leave first a narrow column and then one wide space for the names of the pupils, and thereafter spaces of equal width

1	2	3	4	5	6	7	8	9	10	11	12	13	14
No.	Pu.	CA	M.Sc.	Re. Sc.	Ar. Sc.	Sp. Sc.	T_i	T_r	T_a	T_s	T_e	G_i	G_r

15	16	17	18	19	20	21	22	23	24	25	26	27
G_a	G_s	G_e	G_e Rk.	Tr. Rk.	G_t	G_p	B_i	B_e	B_p	F_e	St. Cl.	Ac. Cl.

Fig. 81. Column headings for a classification-data sheet

sufficient to include numbers of two or three figures. Beginning at the left-hand side with the first narrow column, the columns should be numbered from 1 to 27 in the horizontal line at the top of the page, the wide column being numbered 2. When this has been completed, the label of the classification-data sheet can be placed in the margin at the top of the page. This label should include the title of the sheet, space for the name of the school, the grade, the name of the teacher, and the date. When the ruling and numbering have been completed, the headings will appear as in Fig. 81.

Step 4. *Placing pupils' names in alphabetical order and numbering them*

The horizontal line under the numbers at the top of the classification-data sheet should be reserved for the designation of the columns. The column numbered 1 should be used for numbering the pupils; it can therefore be labeled

"No.," and each line beneath it be numbered from 1 to the number of pupils in the class. When this has been done, the names of the pupils should be placed in alphabetical order in the second column, which can be labeled "Pu." When this has been done, the classification-data sheet for the fifth grade, which has been considered, will appear as shown in columns 1 and 2 of Fig. 82.

The classification-data sheet for the first grade, which has been used in illustration, will need fewer columns, since only two tests were used. In practice it is wise to have a standard classification sheet mimeographed for use, as frequently several copies will be desired. In this case the data for the first grade can be placed on such a sheet, leaving a few columns blank, as will be shown later.

STEP 5. *Filling in chronological ages*

Column 3 is reserved for the chronological ages of the pupils. These data are obtained as a rule from the intelligence-test data, where they are usually required with greater accuracy than with other tests. In the cases here considered the chronological ages are given in years and months to the last full completed month of age. As was explained previously, the error in this is slight, and the calculation of the ages is much less tedious than might otherwise be the case. The column numbered 3 is labeled "CA," to indicate chronological age, and the ages of the various pupils when filled in will have the appearance of column 3 of Fig. 82.

STEP 6. *Transferring raw scores and T scores to the sheet*

This step in the procedure is to copy in their appropriate spaces the raw scores and T scores for the various tests which were obtained in Steps 1 and 2. In the space immediately below the column numbered 4 should be entered the label "M.Sc.," indicating "raw score, mental test," and each space below should contain the raw score in the mental test

received by each child named in the appropriate horizontal line. These data should be taken from the sheet prepared in Step 1.

Column 5 should be labeled "Re. Sc.," indicating "raw score in the reading test." These data should be taken from the sheet given in Table LXX for the first grade and from Table LXXI for the fifth grade.

Column 6 should be labeled "Ar. Sc.," indicating "raw scores in the arithmetic test," and the data for the fifth grade can be taken from Table LXXII. There was no arithmetic test given to the first grade; so this column can be left blank in the classification-data sheet for that grade.

Column 7 should be labeled "Sp. Sc." to indicate "raw scores in the spelling test," and the data for the fifth grade will be found in Table LXXIII. The first grade was not given a test in spelling, so this column will be left blank.

The next four columns should be labeled for the various T scores which were found in Steps 1 and 2. In practice these scores can be filled in at the time that the raw scores are filled in, unless the examiner waits until the raw-score data are placed on the classification sheets before determining the T scores. Column 8 is labeled "T_i," indicating the T scores in intelligence; column 9 "T_r," indicating the T scores in reading; column 10 "T_a," indicating the T scores in arithmetic; and column 11 "T_s," indicating the T scores in spelling. When the columns have been labeled, the various scores received by the pupils can be placed in their appropriate spaces. These scores are shown in Fig. 82 for the fifth grade and in Fig. 83 for the first grade.

STEP 7. *Finding the average T score*

The purpose of the average T score is to combine the T scores received by any pupil into one score which will give a general indication of the level of his achievement level. In

1	2	3	4	5	6	7	8	9	10	11	12	13	14
No.	Pu.	CA	M. Sc.	Re. Sc.	Ar. Sc.	Sp. Sc.	T_i	T_r	T_a	T_s	T_e	G_i	G_r
1	A	11–5	72.0	17	16	25	39	42	34	43	40		
2	B	11–9	106.3	20	21	33	49	47	41	49	45		
3	C	10–5	119.3	26	26	41	53	60	50	57	56		
4	D	12–2	100.3	18	24	29	47	43	46	46	45		
5	E	12–0	66.9	16	24	22	38	40	46	40	42		
6	F	13–3	115.2	18	21	18	52	43	41	37	41		
7	G	10–10	145.5	25	24	37	62	58	46	53	53		
8	H	11–10	124.4	23	33	35	55	53	70	51	58		
9	I	12–2	78.1	20	25	22	41	47	48	40	46		
10	J	14–6	109.8	23	25	27	50	53	48	44	50		
11	K	13–0	127.8	21	22	30	55	49	42	47	46		
12	L	10–5	120.5	23	26	36	53	53	50	52	52		
13	M	11–7	128.7	22	21	36	56	51	41	52	48		
14	N	13–6	77.8	14	16	29	41	38	34	46	38		
15	O	12–0	114.0	16	21	22	52	40	41	40	40		
16	P	12–10	82.8	16	22	23	42	40	42	41	41		
17	Q	11–5	107.4	18	25	33	49	43	48	49	46		
18	R	11–3	117.8	23	30	32	52	53	59	48	54		
19	S	10–9	147.6	28	24	47	63	67	46	67	60		
20	T	11–6	93.0	14	23	27	45	38	44	44	41		
21	U	10–8	124.9	25	26	43	55	58	50	60	56		
22	V	11–10	90.3	12	21	31	44	36	41	48	40		
23	W	11–10	93.4	20	23	29	45	47	44	46	46		
24	X	10–1	133.8	25	28	33	57	58	54	49	55		
25	Y	10–9	114.1	23	28	29	52	53	54	46	52		
26	Z	10–9	141.7	23	21	31	60	53	41	48	48		
27	AA	11–2	79.6	12	21	37	41	36	41	53	41		
28	BB	11–1	95.1	16	21	26	46	40	41	44	41		
29	CC	10–3	154.4	23	22	31	66	53	42	48	49		
30	DD	11–2	100.8	22	29	35	47	51	56	51	53		
31	EE	10–7	96.4	22	14	28	46	51	32	45	44		
32	FF	12–5	87.6	16	27	28	43	40	51	45	45		
33	GG	11–3	75.0	17	26	35	40	42	50	51	46		
34	HH	11–0	130.0	22	27	30	56	51	51	47	50		
35	II	11–11	79.4	16	21	30	41	40	41	47	42		
36	JJ	13–0	84.7	18	19	25	43	43	38	43	41		
37	KK	10–8	91.9	21	24	32	44	49	46	48	48		

FIG. 82. Classification-data sheet containing raw T scores for pupils of fifth grade

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
No.	Pu.	CA	M. Sc.	Re. Sc.	Ar. Sc.	Sp. Sc.	T_i	T_r	T_a	T_s	T_e	G_i	G_r	G_a	G_s	G_e	G_e Rk.	T_R Rk.	G_t
1	A	6–10	34	6			16	12			12								
2	B	6–3	41	17			24	23			23								
3	C	6–10	33	2			15	8			8								
4	D	6–11	28	13			8	19			19								
5	E	6–4	17	2			− 5	8			8								
6	F	6–11	47	21			29	27			27								
7	G	7–3	37	7			20	13			13								
8	H	8–6	11	3			−12	9			9								
9	I	6–9	22	5			− 2	11			11								
10	J	6–8	40	11			23	17			17								
11	K	6–2	38	5			20	11			11								
12	L	7–0	40	26			23	31			31								
13	M	6–5	27	4			6	10			10								
14	N	6–6	34	2			16	8			8								
15	O	7–7	37	21			20	27			27								
16	P	6–4	29	17			11	23			23								
17	Q	7–2	32	11			13	17			17								
18	R	7–0	28	2			8	8			8								
19	S	6–10	22	2			− 2	8			8								
20	T	6–7	22	4			− 2	10			10								
21	U	6–0	36	16			19	22			22								
22	V	6–11	23	4			0	10			10								
23	W	6–5	29	17			9	23			23								
24	X	6–5	35	7			17	13			13								
25	Y	7–2	30	19			11	25			25								
26	Z	6–10	29	5			9	11			11								
27	AA	7–5	39	8			22	14			14								
28	BB	7–9	25	13			2	19			19								
29	CC	7–2	46	28			28	33			33								
30	DD	7–4	51	31			32	36			36								

FIG. 83. Classification-data sheet for first grade (inclusion of raw and T scores)

order to have the T score show this general level, it is wise to weight the relative value of these tests and find a weighted average. In this particular situation it was believed that reading as an ability was the most important asset of a pupil and was more nearly indicative of the achievement level in those subjects which involved reading and which were not measured by any of the tests, and it was also believed that the arithmetic test was more valuable than the spelling test. The reading test was then weighted as of three times the value of the spelling as a general achievement value and twice the value of the arithmetic. The following formula was therefore used:

$$\frac{3\,T_r + 2\,T_a + T_s}{6} = T_e.$$

The symbol T_e indicates the weighted average T score for each pupil as determined by the foregoing formula, in the case of the pupils in the fifth grade. In the first grade, where only one test was given, the T score in reading was taken as the T_e and used throughout. Other tests might have been given in the first grade but were thought to be unnecessary.

Applying this formula to the cases of the pupils in the fifth grade, it will be seen that Pupil 1 has a T_r of 42, a T_a of 34, and a T_s of 43:

$$\frac{3(42) + 2(34) + 43}{6} = 40\ (T_e).$$

Pupil 2 has a T_r of 47, a T_a of 41, and a T_s of 49. Making the substitution in the weighting formula given above, the T_e for Pupil 2 becomes 45:

$$\frac{3(47) + 2(41) + 49}{6} = 45.$$

This process should be continued for all the fifth-grade pupils, and the scores added to the classification-data sheet in column 12, as in Fig. 82.

STEP 8. *Determining the grade scores*

Grade scores for each pupil are determined according to the process given in Chapter IX, p. 203.

A grade table should be constructed either from grade or age norms, or possibly from the T table. These grade scores can be obtained for some tests, as was explained earlier, or can easily be made up from the data given in the test manuals according to the methods given in Chapter IX. In the T scores obtained for the National Intelligence Test in the fifth grade, the G scores were obtained from the T-G table given for that test, as reproduced in Table XXXVII. These become G_i scores, or grade scores in intelligence. The grade score in reading (G_r) was obtained from a T-G table constructed for the Thorndike-McCall Reading Scale. The grade score in arithmetic (G_a) was obtained from a T-G table constructed for the Woody-McCall Arithmetic Test. The grade score in spelling (G_s) was obtained from a T-G table constructed for the Morrison-McCall Spelling Scale.[1] These G scores for intelligence (G_i), for reading (G_r), for arithmetic (G_a), and for spelling (G_s) are placed in columns numbered 13, 14, 15, and 16 on the classification-data sheet.

STEP 9. *Determining the average grade score in achievement* (G_e)

To combine the various grade scores in educational achievement into a single score which will give a general indication of the average grade level of a pupil in his educational achievement, it is necessary to find a weighted score similar to the weighted T score previously found. In this case it was believed that the same reasoning holding for the weighting of the T score would hold with respect to the grade levels of the pupils. This was that the grade level in reading had a high correlation with the grade level of the pupil in the

[1] These T-G tables can be obtained from Teachers College Bureau of Publications.

subjects in which reading played a large part, that the grade level in arithmetic was indicative of the level of the pupil in school skills, and that the grade level in spelling was valuable, but less valuable than either of the other achievement measures. Thus the formula used for the weighted average was as follows:

$$\frac{3\,G_r + 2\,G_a + G_s}{6} = G_e.$$

Applying this formula to the scores obtained by the first pupil for the fifth grade, whose G_r was 4.9, his G_a, 4.7, and his G_s, 4.7, the substitution becomes

$$\frac{3(4.9) + 2(4.7) + 4.7}{6} = 4.8.$$

For the second pupil the substitution given below results in a G_e of 5.9. This pupil has a G_r of 5.8, a G_a of 5.8, and a G_s of 6.2. The substitution becomes

$$\frac{3(5.8) + 2(5.8) + 6.2}{6} = 5.9.$$

Similar procedure should be carried out for all the pupils in the class, giving for all of them a weighted average score in general educational achievement. The first grade (Fig. 85). having had but one test, a reading test, does not have a weighted average score, and the G_e; is in reality the grade score in the reading test. Column 17 is reserved for the G_e; scores. When this step has been carried out, the classification-data sheet will appear as in Fig. 84 for the fifth grade, and as in Fig. 85 for the first grade.

STEP 10. *Determining the teacher's estimate of the pupil's worthiness for promotion*

The average scores which have now been obtained are those of achievement, as represented by the weighted scores in educational achievement (G_e), and a grade score in intelligence or ability, as evidenced by the G_i. Classification might

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
No.	Pu.	CA	M. Sc.	Rr. Sc.	Ar. Sc.	Sp. Sc.	T_i	T_r	T_a	T_s	T_e	G_i	G_r	G_a	G_s	G_e	G_e Rk.	Tr Rk.
1	A	11-5	72.0	17	16	25	39	42	34	43	40	4.4	4.9	4.7	4.7	4.8		
2	B	11-9	106.3	20	21	33	49	47	41	49	45	6.4	5.8	5.8	6.2	5.9		
3	C	10-5	119.3	26	26	41	53	60	50	57	56	7.2	8.8	7.2	8.0	8.1		
4	D	12-2	100.3	18	24	29	47	43	46	46	45	6.0	5.1	6.5	5.4	5.6		
5	E	12-0	66.9	16	24	22	38	40	46	40	42	4.2	4.5	6.5	4.2	5.1		
6	F	13-3	115.2	18	21	18	52	43	41	37	41	7.0	5.1	5.8	3.5	5.1		
7	G	10-10	145.5	25	24	37	62	58	46	53	53	9.0	8.0	6.5	7.0	7.3		
8	H	11-10	124.4	23	33	35	55	53	70	51	58	7.6	6.8	11.6	6.6	8.4		
9	I	12-2	78.1	20	25	22	41	47	48	40	46	4.8	5.8	6.8	4.2	5.9		
10	J	14-6	109.8	23	25	27	50	53	48	44	50	6.6	6.8	6.8	5.1	6.5		
11	K	13-0	127.8	21	22	30	55	49	42	47	46	7.6	6.1	6.0	5.6	6.0		
12	L	10-5	120.5	23	26	36	53	53	50	52	52	7.2	6.8	7.2	6.8	6.9		
13	M	11-7	128.7	22	21	36	56	51	41	52	48	7.8	6.5	5.8	6.8	6.3		
14	N	13-6	77.8	14	16	29	41	38	34	46	38	4.8	4.2	4.7	5.4	4.6		
15	O	12-0	114.0	16	21	22	52	40	41	40	40	7.0	4.5	5.8	4.2	4.9		
16	P	12-10	82.8	16	22	23	42	40	42	41	41	5.0	4.5	6.0	4.3	5.0		
17	Q	11-5	107.4	18	25	33	49	43	48	49	46	6.4	5.1	6.8	6.2	5.9		
18	R	11-3	117.8	23	30	32	52	53	59	48	54	7.0	6.8	9.3	6.0	7.5		
19	S	10-9	147.6	28	24	47	63	67	46	67	60	9.2	12.6	6.5	11.2	10.3		
20	T	11-6	93.0	14	23	27	45	38	44	44	41	5.6	4.2	6.2	5.1	5.0		
21	U	10-8	124.9	25	26	43	55	58	50	60	56	7.6	8.0	7.2	8.8	7.9		
22	V	11-10	90.3	12	21	31	44	36	41	48	40	5.4	3.8	5.8	5.8	4.8		
23	W	11-10	93.4	20	23	29	45	47	44	46	46	5.6	5.8	6.2	5.4	5.9		
24	X	10-1	133.8	25	28	33	57	58	54	49	55	8.0	8.0	8.0	6.2	7.7		
25	Y	10-9	114.1	23	28	29	52	53	54	46	52	7.0	6.8	8.0	5.4	7.0		
26	Z	10-9	141.7	23	21	31	60	53	41	48	48	8.6	6.8	5.8	5.8	6.3		
27	AA	11-2	79.6	12	21	37	41	36	41	53	41	4.8	3.8	5.8	7.0	5.0		
28	BB	11-1	95.1	16	21	26	46	40	41	44	41	5.8	4.5	5.8	4.9	5.0		
29	CC	10-8	154.4	23	22	31	66	53	42	48	49	9.7	6.8	6.0	5.8	6.4		
30	DD	11-2	100.8	22	29	35	47	51	56	51	53	6.0	6.5	8.6	6.6	7.2		
31	EE	10-7	96.4	22	14	28	46	51	32	45	44	5.8	6.5	4.2	5.2	5.5		
32	FF	12-5	87.6	16	27	28	43	40	51	45	45	5.2	4.5	7.5	5.2	5.6		
33	GG	11-3	75.0	17	26	35	40	42	50	51	46	4.6	4.9	7.2	6.6	6.0		
34	HH	11-0	130.0	22	27	30	56	51	51	47	50	7.8	6.5	7.5	5.6	6.7		
35	II	11-11	79.4	16	21	30	41	40	41	47	42	4.8	4.5	5.8	5.6	5.1		
36	JJ	8-0	84.7	18	19	25	43	43	38	43	41	5.2	5.1	5.3	4.7	5.1		
37	KK	10-8	91.9	21	24	32	44	49	46	48	48	5.4	6.1	6.5	6.0	6.2		

FIG. 84. Classification-data sheet for fifth grade (transmutation of T to G scores)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
No.	Pu.	CA	M. Sc.	Re. Sc.	Ar. Sc.	Sp. Sc.	T_i	T_r	T_a	T_s	T_e	G_i	G_r	G_a	G_s	G_e	G_e Rk.	Tr. Rk.	G_t
1	A	6–10	34	6			16	12			12	1.7	1.8			1.8			
2	B	6–3	41	17			24	23			23	2.5	2.5			2.5			
3	C	6–10	33	2			15	8			8	1.6	1.6			1.6			
4	D	6–11	28	13			8	19			19	1.1	2.2			2.2			
5	E	6–4	17	2			−5	8			8	0.5	1.6			1.6			
6	F	6–11	47	21			29	27			27	3.6	2.9			2.9			
7	G	7–3	37	7			20	13			13	2.0	1.8			1.8			
8	H	8–6	11	3			−12	9			9	0.3	1.6			1.6			
9	I	6–9	22	5			−2	11			11	0.7	1.7			1.7			
10	J	6–8	40	11			23	17			17	2.4	2.1			2.1			
11	K	6–2	38	5			20	11			11	2.2	1.7			1.7			
12	L	7–0	40	26			23	31			31	2.4	3.4			3.4			
13	M	6–5	27	4			6	10			10	1.0	1.7			1.7			
14	N	6–6	34	2			16	8			8	1.7	1.6			1.6			
15	O	7–7	37	21			20	27			27	2.2	2.9			2.9			
16	P	6–4	29	17			11	23			23	1.3	2.5			2.5			
17	Q	7–2	32	11			13	17			17	1.5	2.1			2.1			
18	R	7–0	28	2			8	8			8	1.1	1.6			1.6			
19	S	6–10	22	2			−2	8			8	0.7	1.6			1.6			
20	T	6–7	22	4			−2	10			10	0.7	1.7			1.7			
21	U	6–0	36	16			19	22			22	1.9	2.5			2.5			
22	V	6–11	23	4			0	10			10	0.7	1.7			1.7			
23	W	6–5	29	17			9	23			23	1.2	2.5			2.5			
24	X	6–5	35	7			17	13			13	1.8	1.8			1.8			
25	Y	7–2	30	19			11	25			25	1.3	2.7			2.7			
26	Z	6–10	29	5			9	11			11	1.2	2.1			2.1			
27	AA	7–5	39	8			22	14			14	2.3	1.9			1.9			
28	BB	7–9	25	13			2	19			19	0.8	2.2			2.2			
29	CC	7–2	46	28			28	33			33	3.5	3.6			3.6			
30	DD	7–4	51	31			32	36			36	4.3	3.9			3.9			

FIG. 85. Classification-data sheet for first grade (inclusion of grade scores)

be made on these two grounds, since they represent two of the more important items in the equipment of a pupil for the determination of his place in school. Although they tell what he has accomplished in the way of learning and also with what success he is likely to continue his learning, there are several other factors concerning his success which are neither measured by these two units nor included in them. Such factors as the emotional equipment of the child; his general attitude toward school work, his teachers, other pupils, and the like; his ability to respond to leadership or to accept responsibility; his ability to utilize his knowledge, to organize it, or to adapt it to his purposes, — such factors as these are all important in determining a pupil's status and his ability to cope with school work. The one person who is most familiar with a pupil, who has watched his methods of work, who has aided him in his difficulties, who has seen the results of his attitudes, who is familiar with his emotional reaction to the school, and who is acquainted with his general maturity is the classroom teacher. It is then reasonable to expect that the classroom teacher, if asked to rate the pupils for promotion purposes, would necessarily take into account the factors aside from achievement and ability which should be considered. It is neither expected nor necessary that the teacher's knowledge of the pupils' achievement should be disregarded.

To determine a pupil's grade status in terms of a teacher's judgment (G_t), therefore, it is necessary for the teacher to rate the pupils from best to poorest as he would recommend them for promotion. Then since this judgment represents the respective placement of the pupils, and since the average grade score in achievement represents their respective placement in attainment, the grade ranking of the pupil in terms of a teacher's judgment should be the grade score equivalent to the respective ranking within the average grade scores in achievement.

The technic for transmuting the ranks in teachers' judgments of pupils into grade scores in teachers' judgments (G_t) is as follows:

List first, in column 18, the rank order of the grade scores in achievement (G_e).[1] An example of this, with five scores, is shown in Table LXXIV. The first column indicates the pupils by letter; the second shows the G_e scores of the five pupils; and the third shows their rank order, the best score being ranked as 1 and the poorest score as 5.

TABLE LXXIV. DETERMINATION OF G_e RANKINGS FROM G_e SCORES FOR FIVE PUPILS

PUPIL	G_e	G_e RK.
A	5.3	4
B	5.9	2
C	6.1	1
D	5.7	3
E	5.0	5

The pupil with the G_e score of 6.1 has the highest G_e, and his G_e rank is 1; the pupil with the score of 5.9 has the next highest G_e, and his rank is therefore 2. The remaining ranks are as shown in Table LXXIV.

When this column has been completed, the ranking of the pupils for promotion, as made by the teacher independently of any tests and showing his best judgment of the fitness of the pupils for promotion, should be placed in a parallel column. For this group of five pupils let it be assumed that this ranking is as given in Table LXXV, which shows that the teacher thought Pupil B most fit for promotion and gave him a ranking of 1, Pupil C next most fit with a ranking of 2, and so on. Adding this column of ranks (Tr. Rk.) to Table LXXIV gives Table LXXV.

TABLE LXXV. INCLUSION OF TEACHERS' RANKINGS FOR THE FIVE PUPILS OF TABLE LXXIV

PUPIL	G_e	G_e RK.	TR. RK.
A	5.3	4	3
B	5.9	2	1
C	6.1	1	2
D	5.7	3	5
E	5.0	5	4

[1] See Chapter VII, p. 146.

The third element in this transmutation is to give to the various pupils the grade scores equivalent to their teacher's ranking, and to call these scores the teacher's grade scores (G_t). The transmutation is followed through in Table LXXVI by means of arrows. Pupil B, with a Tr. Rk. of 1, is given a G_t score equivalent to the highest G_e score, which is 6.1 and which is ranked 1; therefore, 6.1 is the G_t score for Pupil B. The other scores are found in the same way and are indicated in the last column of Table LXXVI.

TABLE LXXVI. TRANSMUTATION OF G_e TO G_t

Pupil	G_e	G_eRk.	Tr.Rk.	G_t
A	5.3	4	3	5.7
B	5.9	2	1	6.1
C	6.1	1	2	5.9
D	5.7	3	5	5.0
E	5.0	5	4	5.3

The same procedure, in more extended form because of the larger number of pupils involved, should be carried out on the classification-data sheets. In the example which has been used, column 18 is reserved for the rankings of the G_e scores (G_eRk.), column 19 for the teacher's rankings (Tr. Rk.), and column 20 for the teacher's grade scores (G_t). In practice it is unnecessary to use the arrows as shown in Table LXXVI, where they were used merely to show the derivation of the transmuted scores. Probably the easiest method is to begin with the first pupil on the sheet, determine his teacher's ranking as found in column 19, find the equivalent ranking in the G_e-ranking column, find the G_e score for that ranking, and place it in the appropriate space in column 20. Thus Pupil 1 in the fifth grade has a Tr. Rk. of 33. The pupil with the G_e ranking of 33 has a G_e of 5.0, which is then placed in column 20 as a G_t score for Pupil 1.

Pupil 2 has a Tr. Rk. of 24. The pupil with a G_eRk. of 24 has a G_e score of 5.6, which is then transferred to column 20 and entered there as the G_t score for Pupil 2. This pro-

cedure, carried out for all the pupils, gives a G_t rating for each of them. Fig. 86 shows the classification-data sheet for the fifth grade previously used for illustration, with the transmuted G_t scores placed on the sheet in column 20 and with the two sets of rankings in columns 18 and 19.

Fig. 87 shows the transmuted ratings for G_t for the first grade previously used for illustration. Here it will be noted that the G_e score is the same as the G_r score, since only one achievement test was given to the grade.

STEP 11. *Determining a general promotion rating*

In the previous steps which have been completed three different general-average grade scores have been determined. Transmuting the raw score into a T score and from that into a grade score in intelligence from the data found in the intelligence test produced a rating of ability in terms of grade level; transmuting a series of raw scores in achievement tests into T scores in achievement, transmuting the T scores into grade scores in achievement, and then combining these grade scores into one grade score in education produced a single score for each pupil in achievement; finally, transmuting the rating of the teacher in terms of the worthiness for promotion of the pupils into a grade score of the teacher produced a single score containing many elements not otherwise measured. To combine these three general scores constituting ratings of ability, achievement, and worthiness will give a general rating for promotion purposes. This score may be called the grade status for promotion and be symbolized by G_p.

A consideration of these three scores will show that though they should all be taken into consideration for promotion purposes, they should not have equal weight in making the determination. The score of ability was obtained from a highly objective test which was carefully given, carefully

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
No.	Pu.	CA	M. Sc.	RE. Sc.	AR. Sc.	SP. Sc.	T_i	T_r	T_a	T_s	T_e	G_i	G_r	G_a	G_s	G_e	G_e RK.	TR. RK.	G_t	G_p	Bi	B_e
1	A	11–5	72.0	17	16	25	39	42	34	43	40	4.4	4.9	4.7	4.7	4.8	36	33	5.0			
2	B	11–9	106.3	20	21	33	49	47	41	49	45	6.4	5.8	5.8	6.2	5.9	21	24	5.6			
3	C	10–5	119.3	26	26	41	53	60	50	57	56	7.2	8.8	7.2	8.0	8.1	3	10	6.9			
4	D	12–2	100.3	18	24	29	47	43	46	46	45	6.0	5.1	6.5	5.4	5.6	23	28	5.1			
5	E	12–0	66.9	16	24	22	38	40	46	40	42	4.2	4.5	6.5	4.2	5.1	26	32	5.0			
6	F	13–3	115.2	18	21	18	52	43	41	37	41	7.0	5.1	5.8	3.5	5.1	27	34	4.9			
7	G	10–10	145.5	25	24	37	62	58	46	53	53	9.0	8.0	6.5	7.0	7.3	7	4	7.9			
8	H	11–10	124.4	23	33	35	55	53	70	51	58	7.6	6.8	11.6	6.6	8.4	2	14	6.3			
9	I	12–2	78.1	20	25	22	41	47	48	40	46	4.8	5.8	6.8	4.2	5.9	22	37	4.6			
10	J	14–6	109.8	23	25	27	50	53	48	44	50	6.6	6.8	6.8	5.1	6.5	12	15	6.3			
11	K	13–0	127.8	21	22	30	55	49	42	47	46	7.6	6.1	6.0	5.6	6.0	17	19	5.9			
12	L	10–5	120.5	23	26	36	53	53	50	52	52	7.2	6.8	7.2	6.8	6.9	10	9	7.0			
13	M	11–7	128.7	22	21	36	56	51	41	52	48	7.8	6.5	5.8	6.8	6.3	15	7	7.3			
14	N	13–6	77.8	14	16	29	41	38	34	46	38	4.8	4.2	4.7	5.4	4.6	37	36	4.8			
15	O	12–0	114.0	16	21	22	52	40	41	40	40	7.0	4.5	5.8	4.2	4.9	34	31	5.0			
16	P	12–10	82.8	16	22	23	42	40	42	41	41	5.0	4.5	6.0	4.3	5.0	30	35	4.8			
17	Q	11–5	107.4	18	25	33	49	43	48	49	46	6.4	5.1	6.8	6.2	5.9	19	18	6.0			
18	R	11–3	117.8	23	30	32	52	53	59	48	54	7.0	6.8	9.3	6.0	7.5	6	5	7.7			
19	S	10–9	147.6	28	24	47	63	67	46	67	60	9.2	12.6	6.5	11.2	10.3	1	1	10.3			
20	T	11–6	93.0	14	23	27	45	38	44	44	41	5.6	4.2	6.2	5.1	5.0	31	29	5.1			
21	U	10–8	124.9	25	26	43	55	58	50	60	56	7.6	8.0	7.2	8.8	7.9	4	3	8.1			
22	V	11–10	90.3	12	21	31	44	36	41	48	40	5.4	3.8	5.8	5.8	4.8	35	20	5.9			
23	W	11–10	93.4	20	23	29	45	47	44	46	46	5.6	5.8	6.2	5.4	5.9	20	23	5.6			
24	X	10–1	133.8	25	28	33	57	58	54	49	55	8.0	8.0	8.0	6.2	7.7	5	2	8.4			
25	Y	10–9	114.1	23	28	29	52	53	54	46	52	7.0	6.8	8.0	5.4	7.0	9	13	6.4			
26	Z	10–9	141.7	23	21	31	60	53	41	48	48	8.6	6.8	5.8	5.8	6.3	14	8	7.2			
27	AA	11–2	79.6	12	21	37	41	36	41	53	41	4.8	3.8	5.8	7.0	5.0	32	26	5.1			
28	BB	11–1	95.1	16	21	26	46	40	41	44	41	5.8	4.5	5.8	4.9	5.0	33	12	6.5			
29	CC	10–3	154.4	23	22	31	66	53	42	48	49	9.7	6.8	6.0	5.8	6.4	13	11	6.7			
30	DD	11–2	100.8	22	29	35	47	51	56	51	53	6.0	6.5	8.6	6.6	7.2	8	6	7.5			
31	EE	10–7	96.4	22	14	28	46	51	32	45	44	5.8	6.5	4.2	5.2	5.5	25	30	5.5			
32	FF	12–5	87.6	16	27	28	43	40	51	45	45	5.2	4.5	7.5	5.2	5.6	24	17	6.0			
33	GG	11–3	75.0	17	26	35	40	42	50	51	46	4.6	4.9	7.2	6.6	6.0	18	22	5.9			
34	HH	11–0	130.0	22	27	30	56	51	51	47	50	7.8	6.5	7.5	5.6	6.7	11	16	6.2			
35	II	11–11	79.4	16	21	30	41	40	41	47	42	4.8	4.5	5.8	5.6	5.1	28	27	5.1			
36	JJ	13–0	84.7	18	19	25	43	43	38	43	41	5.2	5.1	5.3	4.7	5.1	29	21	5.9			
37	KK	10–8	91.9	21	24	32	44	49	46	48	48	5.4	6.1	6.5	6.0	6.2	16	25	5.5			

FIG. 86. Classification-data sheet for fifth grade (grade status according to the teacher)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
No.	Pu.	CA	M.Sc.	Re. Sc.	Ar. Sc.	Sp. Sc.	T_i	T_r	T_a	T_s	T_e	G_i	G_r	G_a	G_s	G_e	G_eR_K.	Tr. R_K.	G_t
1	A	6–10	34	6			16	12			12	1.7	1.8			1.8	18	20	1.7
2	B	6–3	41	17			24	23			23	2.5	2.5			2.5	10	11	2.2
3	C	6–10	33	2			15	8			8	1.6	1.6			1.6	27	28	1.6
4	D	6–11	28	13			8	19			19	1.1	2.2			2.2	12	1	3.9
5	E	6–4	17	2			−5	8			8	0.5	1.6			1.6	28	30	1.6
6	F	6–11	47	21			29	27			27	3.6	2.9			2.9	5	10	2.5
7	G	7–3	37	7			20	13			13	2.0	1.8			1.8	19	9	2.5
8	H	8–6	11	3			−12	9			9	0.3	1.6			1.6	29	29	1.6
9	I	6–9	22	5			−2	11			11	0.7	1.7			1.7	22	19	1.8
10	J	6–8	40	11			23	17			17	2.4	2.1			2.1	15	8	2.5
11	K	6–2	38	5			20	11			11	2.2	1.7			1.7	23	23	1.7
12	L	7–0	40	26			23	31			31	2.4	3.4			3.4	3	5	2.9
13	M	6–5	27	4			6	10			10	1.0	1.7			1.7	24	21	1.7
14	N	6–6	34	2			16	8			8	1.7	1.6			1.6	30	25	1.6
15	O	7–7	37	21			20	27			27	2.2	2.9			2.9	4	3	3.4
16	P	6–4	29	17			11	23			23	1.3	2.5			2.5	7	12	2.2
17	Q	7–2	32	11			13	17			17	1.5	2.1			2.1	13	15	2.1
18	R	7–0	28	2			8	8			8	1.1	1.6			1.6	25	26	1.6
19	S	6–10	22	2			−2	8			8	0.7	1.6			1.6	26	22	1.7
20	T	6–7	22	4			−2	10			10	0.7	1.7			1.7	20	24	1.7
21	U	6–0	36	16			19	22			22	1.9	2.5			2.5	8	7	2.5
22	V	6–11	23	4			0	10			10	0.7	1.7			1.7	21	27	1.6
23	W	6–5	29	17			9	23			23	1.2	2.5			2.5	9	13	2.1
24	X	6–5	35	7			17	13			13	1.8	1.8			1.8	17	17	1.8
25	Y	7–2	30	19			11	25			25	1.3	2.7			2.7	6	6	2.7
26	Z	6–10	29	5			9	11			11	1.2	2.1			2.1	14	18	1.8
27	AA	7–5	39	8			22	14			14	2.3	1.9			1.9	16	16	1.9
28	BB	7–9	25	13			2	19			19	0.8	2.2			2.2	11	14	2.1
29	CC	7–2	46	28			28	33			33	3.5	3.6			3.6	2	4	2.9
30	DD	7–4	51	31			32	36			36	4.3	3.9			3.9	1	2	3.6

FIG. 87. Classification-data sheet for first grade (grade status according to the teacher)

scored, and the results carefully treated. The score of achievement was obtained from several achievement tests in which great care was observed in giving, scoring, and interpreting. The teacher's rating, although it was determined with care by the teacher, is undoubtedly of inferior origin as compared with the other scores; and because it is far less objective in its determination, it is far more unreliable in its use. From what has been said, however, it should not be inferred that this score is valueless. As has already been stated, it is a score of value because, crude as it may be, it undoubtedly contains elements of measurement of the pupils which are not contained in any other measure available and which should at least be considered in making a standard for promotion. It is desirable, therefore, to consider this score, but not to give it as great weight in the determination of grade status for promotion as is given to other scores of greater reliability.

In this case, then, the three scores to be used for the determination of grade status for promotion should be weighted. Since the scores for ability and achievement should be weighted more than the score for the teacher's judgment, to allow a weighting of two each for the scores for ability and achievement and one for the teacher's rating will accomplish this purpose. The formula, therefore, would be

$$\frac{2\,(G_i) + 2\,(G_e) + G_t}{5} = G_p.$$

In this case the division is by five instead of six, as in the previous weighting, because there are five scores instead of six which make up the final weighting.

By applying this formula to the cases which have been illustrated, in the fifth grade the first pupil has a G_i of 4.4, a G_e of 4.8, and a G_t of 5.0. This means that he has a fourth-grade status at the end of the fourth month in intelligence, a

fourth-grade status at the end of the eighth month in educational achievement, and has qualities which his teacher believes make him more worthy for promotion than his educational status would indicate. Applying the formula given above, his grade status for promotion would become 4.7, or a grade status equivalent to the fourth grade at the end of the seventh month of school.

$$\frac{2\,(4.4) + 2\,(4.8) + 5.0}{5} = 4.7.$$

The second pupil has a G_i of 6.4, equivalent to the sixth grade at the end of the fourth month of school in intelligence; a G_e of 5.9, equivalent to the fifth grade at the end of the ninth month in educational achievement; and a G_t of 5.6, which shows that the teacher believes him to be about eight months below his intelligence and three months below his educational achievement in his worthiness for promotion. His G_p is 6.0, and is found as follows:

$$\frac{2\,(6.4) + 2\,(5.9) + 5.6}{5} = 6.0.$$

Continuing this process with the scores of all the pupils will result in a list of all the grade levels for promotion, as will be found in Fig. 88, column 21. Because these tests may be given at any time during the year, and because promotions are made either at the beginning or in the middle of the year, the statistical classification of pupils must be determined further. This will be considered later.

For the two classes which have been used for illustration the grade status for promotion was calculated as has been outlined above. The scores for the fifth grade in status for promotion (G_p) are given in column 21 of Fig. 88 and are shown with all the other scores which have been found. The G_p scores for the first grade are added to the scores already found and are included in column 21 of Fig. 89.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
No.	Pu.	CA	M. Sc.	Re. Sc.	Ar. Sc.	Sp. Sc.	T_i	T_r	T_a	T_s	T_e	G_i	G_r	G_a	G_s	G_e	G_e Rk.	Tr. Rk.	G_t	G_p	B_i	B_e	B_p	F_e	St. Cl.	Ac. Cl.
1	A	11-5	72.0	17	16	25	39	42	34	43	40	4.4	4.9	4.7	4.7	4.8	36	33	5.0	4.7						
2	B	11-9	106.3	20	21	33	49	47	41	49	45	6.4	5.8	5.8	6.2	5.9	21	24	5.6	6.0						
3	C	10-5	119.3	26	26	41	53	60	50	57	56	7.2	8.8	7.2	8.0	8.1	3	10	6.9	7.5						
4	D	12-2	100.3	18	24	29	47	43	46	46	45	6.0	5.1	6.5	5.4	5.6	23	28	5.1	5.7						
5	E	12-0	66.9	16	24	22	38	40	46	40	42	4.2	4.5	6.5	4.2	5.1	26	32	5.0	4.7						
6	F	13-3	115.2	18	21	18	52	43	41	37	41	7.0	5.1	5.8	3.5	5.1	27	34	4.9	5.8						
7	G	10-10	145.5	25	24	37	62	58	46	53	53	9.0	8.0	6.5	7.0	7.3	7	4	7.9	8.1						
8	H	11-10	124.4	23	33	35	55	53	70	51	58	7.6	6.8	11.6	6.6	8.4	2	14	6.3	7.7						
9	I	12-2	78.1	20	25	22	41	47	48	40	46	4.8	5.8	6.8	4.2	5.9	22	37	4.6	5.2						
10	J	14-6	109.8	23	25	27	50	53	48	44	50	6.6	6.8	6.8	5.1	6.5	12	15	6.3	6.5						
11	K	13-0	127.8	21	22	30	55	49	42	47	46	7.6	6.1	6.0	5.6	6.0	17	19	5.9	6.6						
12	L	10-5	120.5	23	26	36	53	53	50	52	52	7.2	6.8	7.2	6.8	6.9	10	9	7.0	7.0						
13	M	11-7	128.7	22	21	36	56	51	41	52	48	7.8	6.5	5.8	6.8	6.3	15	7	7.3	7.1						
14	N	13-6	77.8	14	16	29	41	38	34	46	38	4.8	4.2	4.7	5.4	4.6	37	36	4.8	4.7						
15	O	12-0	114.0	16	21	22	52	40	41	40	40	7.0	4.5	5.8	4.2	4.9	34	31	5.0	5.8						
16	P	12-10	82.8	16	22	23	42	40	42	41	41	5.0	4.5	6.0	4.3	5.0	30	35	4.8	5.0						
17	Q	11-5	107.4	18	25	33	49	43	48	49	46	6.4	5.1	6.8	6.2	5.9	19	18	6.0	6.1						
18	R	11-3	117.8	23	30	32	52	53	59	48	54	7.0	6.8	9.3	6.0	7.5	6	5	7.7	7.3						
19	S	10-9	147.6	28	24	47	63	67	46	67	60	9.2	12.6	6.5	11.2	10.3	1	1	10.3	9.9						
20	T	11-6	93.0	14	23	27	45	38	44	44	41	5.6	4.2	6.2	5.1	5.0	31	29	5.1	5.3						
21	U	10-8	124.9	25	26	43	55	58	50	60	56	7.6	8.0	7.2	8.8	7.9	4	3	8.1	7.8						
22	V	11-10	90.3	12	21	31	44	36	41	48	40	5.4	3.8	5.8	5.8	4.8	35	20	5.9	5.3						
23	W	11-10	93.4	20	23	29	45	47	44	46	46	5.6	5.8	6.2	5.4	5.9	20	23	5.6	5.7						
24	X	10-1	133.8	25	28	33	57	58	54	49	55	8.0	8.0	8.0	6.2	7.7	5	2	8.4	8.0						
25	Y	10-9	114.1	23	28	29	52	53	54	46	52	7.0	6.8	8.0	5.4	7.0	9	13	6.4	6.9						
26	Z	10-9	141.7	23	21	31	60	53	41	48	48	8.6	6.8	5.8	5.8	6.3	14	8	7.2	7.4						
27	AA	11-2	79.6	12	21	37	41	36	41	53	41	4.8	3.8	5.8	7.0	5.0	32	26	5.1	4.9						
28	BB	11-1	95.1	16	21	26	46	40	41	44	41	5.8	4.5	5.8	4.9	5.0	33	12	6.5	5.6						
29	CC	10-3	154.4	23	22	31	66	53	42	48	49	9.7	6.8	6.0	5.8	6.4	13	11	6.7	7.8						
30	DD	11-2	100.8	22	29	35	47	51	56	51	53	6.0	6.5	8.6	6.6	7.2	8	6	7.5	6.8						
31	EE	10-7	96.4	22	14	28	46	51	32	45	44	5.8	6.5	4.2	5.2	5.5	25	30	5.0	5.5						
32	FF	12-5	87.6	16	27	28	43	40	51	45	45	5.2	4.5	7.5	5.2	5.6	24	17	6.0	5.5						
33	GG	11-3	75.0	17	26	35	40	42	50	51	46	4.6	4.9	7.2	6.6	6.0	18	22	5.9	5.4						
34	HH	11-0	130.0	22	27	30	56	51	51	47	50	7.8	6.5	7.5	5.6	6.7	11	16	6.2	7.0						
35	II	11-11	79.4	16	21	30	41	40	41	47	42	4.8	4.5	5.8	5.6	5.1	28	27	5.1	5.0						
36	JJ	13-0	84.7	18	19	25	43	43	38	43	41	5.2	5.1	5.3	4.7	5.1	29	21	5.9	5.3						
37	KK	10-8	91.9	21	24	32	44	49	46	48	48	5.4	6.1	6.5	6.0	6.2	16	25	5.5	5.7						

FIG. 88. Classification-data sheet for fifth grade (determination of grade status for promotion)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
No.	Pu.	CA	M Sc.	Re. Sc.	Ar. Sc.	Sp. Sc.	T_i	T_r	T_a	T_s	T_e	G_i	G_r	G_a	G_s	G_e	G_e Rk.	Tr. Rk.	G_t	G_p	B_i	B_e	B_p	F_e	St. Cl.	Ac. Cl.
1	A	6–10	34	6			16	12			12	1.7	1.8			1.8	18	20	1.7	1.7						
2	B	6–3	41	17			24	23			23	2.5	2.5			2.5	10	11	2.2	2.4						
3	C	6–10	33	2			15	8			8	1.6	1.6			1.6	27	28	1.6	1.6						
4	D	6–11	28	13			8	19			19	1.1	2.2			2.2	12	1	3.9	2.1						
5	E	6–4	17	2			−5	8			8	0.5	1.6			1.6	28	30	1.6	1.2						
6	F	6–11	47	21			29	27			27	3.6	2.9			2.9	5	10	2.5	3.1						
7	G	7–3	37	7			20	13			13	2.0	1.8			1.8	19	9	2.5	2.0						
8	H	8–6	11	3			−12	9			9	0.3	1.6			1.6	29	29	1.6	1.1						
9	I	6–9	22	5			−2	11			11	0.7	1.7			1.7	22	19	1.8	1.3						
10	J	6–8	40	11			23	17			17	2.4	2.1			2.1	15	8	2.5	2.3						
11	K	6–2	38	5			20	11			11	2.2	1.7			1.7	23	23	1.7	1.9						
12	L	7–0	40	26			23	31			31	2.4	3.4			3.4	3	5	2.9	2.9						
13	M	6–5	27	4			6	10			10	1.0	1.7			1.7	24	21	1.7	1.4						
14	N	6–6	34	2			16	8			8	1.7	1.6			1.6	30	25	1.6	1.6						
15	O	7–7	37	21			20	27			27	2.2	2.9			2.9	4	3	3.4	2.7						
16	P	6–4	29	17			11	23			23	1.3	2.5			2.5	7	12	2.2	2.0						
17	Q	7–2	32	11			13	17			17	1.5	2.1			2.1	13	15	2.1	1.9						
18	R	7–0	28	2			8	8			8	1.1	1.6			1.6	25	26	1.6	1.4						
19	S	6–10	22	2			−2	8			8	0.7	1.6			1.6	26	22	1.7	1.3						
20	T	6–7	22	4			−2	10			10	0.7	1.7			1.7	20	24	1.7	1.3						
21	U	6–0	36	16			19	22			22	1.9	2.5			2.5	8	7	2.5	2.3						
22	V	6–11	23	4			0	10			10	0.7	1.7			1.7	21	27	1.6	1.3						
23	W	6–5	29	17			9	23			23	1.2	2.5			2.5	9	13	2.1	1.9						
24	X	6–5	35	7			17	13			13	1.8	1.8			1.8	17	17	1.8	1.8						
25	Y	7–2	30	19			11	25			25	1.3	2.7			2.7	6	6	2.7	2.1						
26	Z	6–10	29	5			9	11			11	1.2	2.1			2.1	14	18	1.8	1.7						
27	AA	7–5	39	8			22	14			14	2.3	1.9			1.9	16	16	1.9	2.1						
28	BB	7–9	25	13			2	19			19	0.8	2.2			2.2	11	14	2.1	1.6						
29	CC	7–2	46	28			28	33			33	3.5	3.6			3.6	2	4	2.9	3.4						
30	DD	7–4	51	31			32	36			36	4.3	3.9			3.9	1	2	3.6	4.0						

FIG. 89. Classification-data sheet for first grade (inclusion of G_p scores)

STEP 12. *Determining the B-score for promotion*

The measures which have been determined up to this point have given the status of the pupil without reference to his brightness as a measure of his probable success in promotion, except as it is incidentally included in the judgment of the teacher. Although the measure of intelligence which has been used is a measure of ability, it is not necessarily a measure of rate in learning. For example, of two pupils whose mental ages at the time of testing are equal but whose ages are different, the younger, in order to have reached the same plane of ability in the shorter space of time, has moved at a faster rate than has the older pupil. It is also true that the pupil who has moved at the faster rate, other things being equal, will be capable of continuing to move at a faster rate than the older pupil and therefore at any future time, assuming that both pupils are for the moment equal in mental age, will surpass the older.

In the T scale, as has been outlined in Chapter XI, a means was described whereby B scores could be found for promotion. In determining the worthiness of pupils for promotion, the B score derived from chronological ages is necessary in order to place pupils with those with whom they will be likely to do best — those who are not only of an equal educational status but also of an approximately equal rate of learning.

Two B scores are desired in reclassification: the B score in intelligence (B_i) and the B score in education (B_e). Later chapters will show the use of the B score for other purposes.

The B score in intelligence is found according to the method outlined in Chapter XI. For Pupil 1 the T score in intelligence is 39. The B correction for the chronological age 11–5 is + 4.[1] Therefore the T score, 39, plus the B correction, + 4, gives Pupil 1 a B score in intelligence of 43.

[1] See Table XLVI, p. 275.

Since when two or more scores are to be determined by the use of the same data it is always easier to find the two scores at the same time, and since in this case the B score in education is found by adding to the T scores in education the same B correction as was added to the T scores in intelligence, it is simpler to find these two scores at the same time. In the case of Pupil 1 the T score in education (T_e) is 40. Adding the BCr to this score gives a B_e of 44.

In the case of Pupil 2, the T_i is 49 and the T_e is 45. The BCr for a pupil aged 11–9 is + 3.[1] This gives B scores of 52 for B_i and 48 for B_e.

When the process has been completed for all the pupils, column 22 will be filled with the B_i scores and column 23 with the B_e scores.

The B_i and B_e scores for the fifth grades which have been used for illustration are added to the other data which have been found, and are shown in columns 22 and 23 of Fig. 90.

The B_i and B_e scores for the first-grade pupils added to the other data which have been determined are found in columns 22 and 23 of Fig. 91.

The B score for promotion is logically a combination of the B_i score, which indicates the rate at which a pupil is capable of learning, and the B_e score, which indicates the rate at which he has learned in the past. For example, if a pupil has a relatively high brightness score in ability, as would be indicated by a B_i of more than 50, which is average for pupils of his age, and if he has a relatively low score in educational achievement, which is an indication of the fact that he has not been kept at his optimum speed, it is clearly indicated that he can do better than he has done, but it is not so clearly indicated that he can recover his lost ground without increasing his speed above his optimum rate. His true status for promotion according to his brightness, therefore, would

[1] See Table XLVI, p. 275.

lie somewhere between his status of brightness in intelligence and his status of brightness in achievement.

However, if a pupil possesses a brightness in achievement which exceeds his brightness in ability, it is reasonable to suppose that he has been stimulated more than he should have been in the past, and that his optimum rate (since he is ahead of his proper status in achievement) lies somewhere between his brightness in intelligence and his brightness in achievement.

For any pupil an average of the B_i and B_e scores will show the optimum rate at which he is likely to continue his work. The formula for finding the B score for promotion (B_p) is

$$\frac{B_i + B_e}{2} = B_p.$$

Applying this formula to Pupil 1, whose B_i is 43 and whose B_e is 44, the actual B_p is 43.5. In this and all like cases it is wise to use only whole numbers and to change the decimal to the nearest higher whole number, making the B_p in this case 44.

Applying the formula to the case of Pupil 2, whose B_i is 52 and whose B_e is 48, the B_p is found to be 50.

When the B_p scores have been found they should be entered in column 24. Fig. 90 shows this column added to the classification-data sheet for the fifth grade; Fig. 91, the data added to the classification-data sheet for the first grade.

The interpretation of the B_p is one which will guide a teacher in estimating the desirability of the statistical classification, to be described later. In the case of a child who is young for his grade it will show whether he should be promoted; for the one who is old for his grade, it will show whether he can profitably be promoted so as to be with children nearer his own age. The process of interpretation will be outlined in a later discussion.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
No.	Pu.	CA	M. Sc.	Re. Sc.	Ar. Sc.	Sp. Sc.	T_i	T_r	T_a	T_s	T_e	G_i	G_r	G_a	G_s	G_e	G_e Rk.	Tr. Rk.	G_t	G_p	B_i	B_e	B_p	F_e	St. Cl.	Ac. Cl.
1	A	11–5	72.0	17	16	25	39	42	34	43	40	4.4	4.9	4.7	4.7	4.8	36	33	5.0	4.7	43	44	44			
2	B	11–9	106.3	20	21	33	49	47	41	49	45	6.4	5.8	5.8	6.2	5.9	21	24	5.6	6.0	52	48	50			
3	C	10–5	119.3	26	26	41	53	60	50	57	56	7.2	8.8	7.2	8.0	8.1	3	10	6.9	7.5	62	65	64			
4	D	12–2	100.3	18	24	29	47	43	46	46	45	6.0	5.1	6.5	5.4	5.6	23	28	5.1	5.7	48	46	47			
5	E	12–0	66.9	16	24	22	38	40	46	40	42	4.2	4.5	6.5	4.2	5.1	26	32	5.0	4.7	40	44	42			
6	F	13–3	115.2	18	21	18	52	43	41	37	41	7.0	5.1	5.8	3.5	5.1	27	34	4.9	5.8	49	38	44			
7	G	10–10	145.5	25	24	37	62	58	46	53	53	9.0	8.0	6.5	7.0	7.3	7	4	7.9	8.1	70	61	66			
8	H	11–10	124.4	23	33	35	55	53	70	51	58	7.6	6.8	11.6	6.6	8.4	2	14	6.3	7.7	58	61	60			
9	I	12–2	78.1	20	25	22	41	47	48	40	46	4.8	5.8	6.8	4.2	5.9	22	37	4.6	5.2	42	47	45			
10	J	14–6	109.8	23	25	27	50	53	48	44	50	6.6	6.8	6.8	5.1	6.5	12	15	6.3	6.5	43	43	43			
11	K	13–0	127.8	21	22	30	55	49	42	47	46	7.6	6.1	6.0	5.6	6.0	17	19	5.9	6.6	53	44	49			
12	L	10–5	120.5	23	26	36	53	53	50	52	52	7.2	6.8	7.2	6.8	6.9	10	9	7.0	7.0	62	61	62			
13	M	11–7	128.7	22	21	36	56	51	41	52	48	7.8	6.5	5.8	6.8	6.3	15	7	7.3	7.1	60	52	56			
14	N	13–6	77.8	14	16	29	41	38	34	46	38	4.8	4.2	4.7	5.4	4.6	37	36	4.8	4.7	37	34	36			
15	O	12–0	114.0	16	21	22	52	40	41	40	40	7.0	4.5	5.8	4.2	4.9	34	31	5.0	5.8	54	42	48			
16	P	12–10	82.8	16	22	23	42	40	42	41	41	5.0	4.5	6.0	4.3	5.0	30	35	4.8	5.0	41	40	41			
17	Q	11–5	107.4	18	25	33	49	43	48	49	46	6.4	5.1	6.8	6.2	5.9	19	18	6.0	6.1	53	50	52			
18	R	11–3	117.8	23	30	32	52	53	59	48	54	7.0	6.8	9.3	6.0	7.5	6	5	7.7	7.3	57	59	58			
19	S	10–9	147.6	28	24	47	63	67	46	67	60	9.2	12.6	6.5	11.2	10.3	1	1	10.3	9.9	71	68	70			
20	T	11–6	93.0	14	23	27	45	38	44	44	41	5.6	4.2	6.2	5.1	5.0	31	29	5.1	5.3	49	45	47			
21	U	10–8	124.9	25	26	43	55	58	50	60	56	7.6	8.0	7.2	8.8	7.9	4	3	8.1	7.8	63	64	64			
22	V	11–10	90.3	12	21	31	44	36	41	48	40	5.4	3.8	5.8	5.8	4.8	35	20	5.9	5.3	47	43	45			
23	W	11–10	93.4	20	23	29	45	47	44	46	46	5.6	5.8	6.2	5.4	5.9	20	23	5.6	5.7	48	49	49			
24	X	10–1	133.8	25	28	33	57	58	54	49	55	8.0	8.0	8.0	6.2	7.7	5	2	8.4	8.0	68	66	67			
25	Y	10–9	114.1	23	28	29	52	53	54	46	52	7.0	6.8	8.0	5.4	7.0	9	13	6.4	6.9	60	60	60			
26	Z	10–9	141.7	23	21	31	60	53	41	48	48	8.6	6.8	5.8	5.8	6.3	14	8	7.2	7.4	68	56	62			
27	AA	11–2	79.6	12	21	37	41	36	41	53	41	4.8	3.8	5.8	7.0	5.0	32	26	5.1	4.9	47	47	47			
28	BB	11–1	95.1	16	21	26	46	40	41	44	41	5.8	4.5	5.8	4.9	5.0	33	12	6.5	5.6	52	47	50			
29	CC	10–3	154.4	23	22	31	66	53	42	48	49	9.7	6.8	6.0	5.8	6.4	13	11	6.7	7.8	76	59	68			
30	DD	11–2	100.8	22	29	35	47	51	56	51	53	6.0	6.5	8.6	6.6	7.2	8	6	7.5	6.8	53	59	56			
31	EE	10–7	96.4	22	14	28	46	51	32	45	44	5.8	6.5	4.2	5.2	5.5	25	30	5.5	5.5	54	52	53			
32	FF	12–5	87.6	16	27	28	43	40	51	45	45	5.2	4.5	7.5	5.2	5.6	24	17	6.0	5.5	44	46	45			
33	GG	11–3	75.0	17	26	35	40	42	50	51	46	4.6	4.9	7.2	6.6	6.0	18	22	5.9	5.4	45	51	48			
34	HH	11–0	130.0	22	27	30	56	51	51	47	50	7.8	6.5	7.5	5.6	6.7	11	16	6.2	7.0	63	57	60			
35	II	11–11	79.4	16	21	30	41	40	41	47	42	4.8	4.5	5.8	5.6	5.1	28	27	5.1	5.0	43	44	44			
36	JJ	13–0	84.7	18	19	25	43	43	38	43	41	5.2	5.1	5.3	4.7	5.1	29	21	5.9	5.3	41	39	40			
37	KK	10–8	91.9	21	24	32	44	49	46	48	48	5.4	6.1	6.5	6.0	6.2	16	25	5.5	5.7	52	56	54			

FIG. 90. Classification-data sheet for fifth grade (determination of B scores)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
No.	Pu.	CA	M. Sc.	Re. Sc.	Ar. Sc.	Sp. Sc.	T_i	T_r	T_a	T_s	T_e	G_i	G_r	G_a	G_s	G_e	G_e Rk.	Tr. Rk.	G_t	G_p	B_i	B_e	B_p	F_e	St. Cl.	Ac. Cl.
1	A	6–10	34	6			16	12			12	1.7	1.8			1.8	18	20	1.7	1.7	54	50	52			
2	B	6–3	41	17			24	23			23	2.5	2.5			2.5	10	11	2.2	2.4	69	68	69			
3	C	6–10	33	2			15	8			8	1.6	1.6			1.6	27	28	1.6	1.6	53	46	50			
4	D	6–11	28	13			8	19			19	1.1	2.2			2.2	12	1	3.9	2.1	45	56	51			
5	E	6–4	17	2			−5	8			8	0.5	1.6			1.6	28	30	1.6	1.2	39	52	46			
6	F	6–11	47	21			29	27			27	3.6	2.9			2.9	5	10	2.5	3.1	66	64	65			
7	G	7–3	37	7			20	13			13	2.0	1.8			1.8	19	9	2.5	2.0	53	46	50			
8	H	8–6	11	3			−12	9			9	0.3	1.6			1.6	29	29	1.6	1.1	10	31	21			
9	I	6–9	22	5			−2	11			11	0.7	1.7			1.7	22	19	1.8	1.3	37	50	44			
10	J	6–8	40	11			23	17			17	2.4	2.1			2.1	15	8	2.5	2.3	63	57	60			
11	K	6–2	38	5			20	11			11	2.2	1.7			1.7	23	23	1.7	1.9	67	58	63			
12	L	7–0	40	26			23	31			31	2.4	3.4			3.4	3	5	2.9	2.9	59	67	63			
13	M	6–5	27	4			6	10			10	1.0	1.7			1.7	24	21	1.7	1.4	49	53	51			
14	N	6–6	34	2			16	8			8	1.7	1.6			1.6	30	25	1.6	1.6	58	50	54			
15	O	7–7	37	21			20	27			27	2.2	2.9			2.9	4	3	3.4	2.7	50	57	54			
16	P	6–4	29	17			11	23			23	1.3	2.5			2.5	7	12	2.2	2.0	55	67	61			
17	Q	7–2	32	11			13	17			17	1.5	2.1			2.1	13	15	2.1	1.9	47	51	49			
18	R	7–0	28	2			8	8			8	1.1	1.6			1.6	25	26	1.6	1.4	44	44	44			
19	S	6–10	22	2			−2	8			8	0.7	1.6			1.6	26	22	1.7	1.3	36	46	41			
20	T	6–7	22	4			−2	10			10	0.7	1.7			1.7	20	24	1.7	1.3	39	51	45			
21	U	6–0	36	16			19	22			22	1.9	2.5			2.5	8	7	2.5	2.3	69	72	71			
22	V	6–11	23	4			0	10			10	0.7	1.7			1.7	21	27	1.6	1.3	37	47	42			
23	W	6–5	29	17			9	23			23	1.2	2.5			2.5	9	13	2.1	1.9	52	66	59			
24	X	6–5	35	7			17	13			13	1.8	1.8			1.8	17	17	1.8	1.8	60	56	58			
25	Y	7–2	30	19			11	25			25	1.3	2.7			2.7	6	6	2.7	2.1	45	59	52			
26	Z	6–10	29	5			9	11			11	1.2	2.1			2.1	14	18	1.8	1.7	47	49	48			
27	AA	7–5	39	8			22	14			14	2.3	1.9			1.9	16	16	1.9	2.1	53	45	49			
28	BB	7–9	25	13			2	19			19	0.8	2.2			2.2	11	14	2.1	1.6	30	47	39			
29	CC	7–2	46	28			28	33			33	3.5	3.6			3.6	2	4	2.9	3.4	62	67	65			
30	DD	7–4	51	31			32	36			36	4.3	3.9			3.9	1	2	3.6	4.0	64	68	66			

Fig. 91. Classification-data sheet for first grade (inclusion of B scores)

STEP 13. *Determining the effort score (F)*

Whether or not a pupil can make a success of dealing with the work of a grade or class in which he might be placed because of the evidence as indicated by his grade status for promotion, depends to a large extent upon the effort which he makes. If his effort is such that he is maintaining a normal achievement in terms of his ability, it is probable that his grade status will indicate an optimum placement for him; if, on the other hand, his effort is such that his achievement is less than his ability would warrant, it would be difficult for him to maintain the grade level which his grade status for promotion would indicate for him. It is also true that if a pupil, prodded or pushed or overstimulated, had made an achievement greater than his ability would indicate to be normal and right for him to make, such overachievement would have tended to cause the score for grade status in promotion to show the pupil as deserving of a higher rating than might be best for him.

Therefore, in determining the true status of a pupil it is necessary to take into account his effort rating. This rating, as was stated in Chapter XII, is found by means of the formula

$$(B_e + 50) - B_i = \text{F score.}$$

The effort score for Pupil 1, as found by this formula, is 51, showing that he is slightly higher in effort than might be expected of a pupil of his ability. The substitution becomes

$$(44 + 50) - 43 = 51.$$

For pupil 2, who has a B_e of 48 and a B_i of 52, the substitution becomes

$$(48 + 50) - 52 = 46.$$

On the classification-data sheets column 25 has been reserved for the effort scores as found in this way. Fig. 92 shows the effort scores placed, in addition to the other scores found for the fifth grade; Fig. 93 shows similar data for the first grade.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
No.	Pu.	CA	M. Sc.	Re. Sc.	Ar. Sc.	Sp. Sc.	T_i	T_r	T_a	T_s	T_e	G_i	G_r	G_a	G_s	G_e	G_e Rk.	Tr. Rk.	G_t	G_p	B_i	B_e	B_p	F_e	St. Cl.	Ac. Cl.
1	A	11–5	72.0	17	16	25	39	42	34	43	40	4.4	4.9	4.7	4.7	4.8	36	33	5.0	4.7	43	44	44	51		
2	B	11–9	106.3	20	21	33	49	47	41	49	45	6.4	5.8	5.8	6.2	5.9	21	24	5.6	6.0	52	48	50	46		
3	C	10–5	119.3	26	26	41	53	60	50	57	56	7.2	8.8	7.2	8.0	8.1	3	10	6.9	7.5	62	65	64	53		
4	D	12–2	100.3	18	24	29	47	43	46	46	45	6.0	5.1	6.5	5.4	5.6	23	28	5.1	5.7	48	46	47	48		
5	E	12–0	66.9	16	24	22	38	40	46	40	42	4.2	4.5	6.5	4.2	5.1	26	32	5.0	4.7	40	44	42	54		
6	F	13–3	115.2	18	21	18	52	43	41	37	41	7.0	5.1	5.8	3.5	5.1	27	34	4.9	5.8	49	38	44	39		
7	G	10–10	145.5	25	24	37	62	58	46	53	53	9.0	8.0	6.5	7.0	7.3	7	4	7.9	8.1	70	61	66	41		
8	H	11–10	124.4	23	33	35	55	53	70	51	53	7.6	6.8	11.6	6.6	8.4	2	14	6.3	7.7	58	61	60	53		
9	I	12–2	78.1	20	25	22	41	47	48	40	46	4.8	5.8	6.8	4.2	5.9	22	37	4.6	5.2	42	47	45	55		
10	J	14–6	109.8	23	25	27	50	53	48	44	50	6.6	6.8	6.8	5.1	6.5	12	15	6.3	6.5	43	43	43	50		
11	K	13–0	127.8	21	22	30	55	49	42	47	46	7.6	6.1	6.0	5.6	6.0	17	19	5.9	6.6	53	44	49	41		
12	L	10–5	120.5	23	26	36	53	53	50	52	52	7.2	6.8	7.2	6.8	6.9	10	9	7.0	7.0	62	61	62	49		
13	M	11–7	128.7	22	21	36	56	51	41	52	48	7.8	6.5	5.8	6.8	6.3	15	7	7.3	7.1	60	52	56	42		
14	N	13–6	77.8	14	16	29	41	38	34	46	38	4.8	4.2	4.7	5.4	4.6	37	36	4.8	4.7	37	34	36	47		
15	O	12–0	114.0	16	21	22	52	40	41	40	40	7.0	4.5	5.8	4.2	4.9	34	31	5.0	5.8	54	42	48	38		
16	P	12–10	82.8	16	22	23	42	40	42	41	41	5.0	4.5	6.0	4.3	5.0	30	35	4.8	5.0	41	40	41	49		
17	Q	11–5	107.4	18	25	33	49	43	48	49	46	6.4	5.1	6.8	6.2	5.9	19	18	6.0	6.1	53	50	52	47		
18	R	11–3	117.8	23	30	32	52	53	59	48	54	7.0	6.8	9.3	6.0	7.5	6	5	7.7	7.3	57	59	58	52		
19	S	10–9	147.6	28	24	47	63	67	46	67	60	9.2	12.6	6.5	11.2	10.3	1	1	10.3	9.9	71	68	70	47		
20	T	11–6	93.0	14	23	27	45	38	44	44	41	5.6	4.2	6.2	5.1	5.0	31	29	5.1	5.3	49	45	47	46		
21	U	10–8	124.9	25	26	43	55	58	50	60	56	7.6	8.0	7.2	8.8	7.9	4	3	8.1	7.8	63	64	64	51		
22	V	11–10	90.3	12	21	31	44	36	41	48	40	5.4	3.8	5.8	5.8	4.8	35	20	5.9	5.3	47	43	45	46		
23	W	11–10	93.4	20	23	29	45	47	44	46	46	5.6	5.8	6.2	5.4	5.9	20	23	5.6	5.7	48	49	49	51		
24	X	10–1	133.8	25	28	33	57	58	54	49	55	8.0	8.0	8.0	6.2	7.7	5	2	8.4	8.0	68	66	67	48		
25	Y	10–9	114.1	23	28	29	52	53	54	46	52	7.0	6.8	8.0	5.4	7.0	9	13	6.4	6.9	60	60	60	50		
26	Z	10–9	141.7	23	21	31	60	53	41	48	48	8.6	6.8	5.8	5.8	6.3	14	8	7.2	7.4	68	56	62	38		
27	AA	11–2	79.6	12	21	37	41	36	41	53	41	4.8	3.8	5.8	7.0	5.0	32	26	5.1	4.9	47	47	47	50		
28	BB	11–1	95.1	16	21	26	46	40	41	44	41	5.8	4.5	5.8	4.9	5.0	33	12	6.5	5.6	52	47	50	45		
29	CC	10–3	154.4	23	22	31	66	53	42	48	49	9.7	6.8	6.0	5.8	6.4	13	11	6.7	7.8	76	59	68	33		
30	DD	11–2	100.8	22	29	35	47	51	56	51	53	6.0	6.5	8.6	6.6	7.2	8	6	7.5	6.8	53	59	56	56		
31	EE	10–7	96.4	22	14	28	46	51	32	45	44	5.8	6.5	4.2	5.2	5.5	25	30	5.5	5.5	54	52	53	48		
32	FF	12–5	87.6	16	27	28	43	40	51	45	45	5.2	4.5	7.5	5.2	5.6	24	17	6.0	5.5	44	46	45	52		
33	GG	11–3	75.0	17	26	35	40	42	50	51	46	4.6	4.9	7.2	6.6	6.0	18	22	5.9	5.4	45	51	48	56		
34	HH	11–0	130.0	22	27	30	56	51	51	47	50	7.8	6.5	7.5	5.6	6.7	11	16	6.2	7.0	63	57	60	44		
35	II	11–11	79.4	16	21	30	41	40	41	47	42	4.8	4.5	5.8	5.6	5.1	28	27	5.1	5.0	43	44	44	51		
36	JJ	13–0	84.7	18	19	25	43	43	38	43	41	5.2	5.1	5.3	4.7	5.1	29	21	5.9	5.3	41	39	40	48		
37	KK	10–8	91.9	21	24	32	44	49	46	48	48	5.4	6.1	6.5	6.0	6.2	16	25	5.5	5.7	52	56	54	54		

FIG. 92. Classification-data sheet for fifth grade (determination of effort scores)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
No.	Pu.	CA	M. Sc.	Re. Sc.	Ar. Sc.	Sp. Sc.	T_i	T_r	T_a	T_s	T_e	G_i	G_r	G_a	G_s	G_e	G_e Rk.	Tr. Rk.	G_t	G_p	B_i	B_e	B_p	F_e	Ac. Cl.	St. Cl.
1	A	6–10	34	6			16	12			12	1.7	1.8			1.8	18	20	1.7	1.7	54	50	52	46		
2	B	6–3	41	17			24	23			23	2.5	2.5			2.5	10	11	2.2	2.4	69	68	69	49		
3	C	6–10	33	2			15	8			8	1.6	1.6			1.6	27	28	1.6	1.6	53	46	50	43		
4	D	6–11	28	13			8	19			19	1.1	2.2			2.2	12	1	3.9	2.1	45	56	51	61		
5	E	6–4	17	2			−5	8			8	0.5	1.6			1.6	28	30	1.6	1.2	39	52	46	63		
6	F	6–11	47	21			29	27			27	3.6	2.9			2.9	5	10	2.5	3.1	66	64	65	48		
7	G	7–3	37	7			20	13			13	2.0	1.8			1.8	19	9	2.5	2.0	53	46	50	43		
8	H	8–6	11	3			−12	9			9	0.3	1.6			1.6	29	29	1.6	1.1	10	31	21	71		
9	I	6–9	22	5			−2	11			11	0.7	1.7			1.7	22	19	1.8	1.3	37	50	44	63		
10	J	6–8	40	11			23	17			17	2.4	2.1			2.1	15	8	2.5	2.3	63	57	60	44		
11	K	6–2	38	5			20	11			11	2.2	1.7			1.7	23	23	1.7	1.9	67	58	63	41		
12	L	7–0	40	26			23	31			31	2.4	3.4			3.4	3	5	2.9	2.9	59	67	63	58		
13	M	6–5	27	4			6	10			10	1.0	1.7			1.7	24	21	1.7	1.4	49	53	51	54		
14	N	6–6	34	2			16	8			8	1.7	1.6			1.6	30	25	1.6	1.6	58	50	54	42		
15	O	7–7	37	21			20	27			27	2.2	2.9			2.9	4	3	3.4	2.7	50	57	54	57		
16	P	6–4	29	17			11	23			23	1.3	2.5			2.5	7	12	2.2	2.0	55	67	61	62		
17	Q	7–2	32	11			13	17			17	1.5	2.1			2.1	13	15	2.1	1.9	47	51	49	54		
18	R	7–0	28	2			8	8			8	1.1	1.6			1.6	25	26	1.6	1.4	44	44	44	50		
19	S	6–10	22	2			−2	8			8	0.7	1.6			1.6	26	22	1.7	1.3	36	46	41	60		
20	T	6–7	22	4			−2	10			10	0.7	1.7			1.7	20	24	1.7	1.3	39	51	45	62		
21	U	6–0	36	16			19	22			22	1.9	2.5			2.5	8	7	2.5	2.3	69	72	71	53		
22	V	6–11	23	4			0	10			10	0.7	1.7			1.7	21	27	1.6	1.3	37	47	42	60		
23	W	6–5	29	17			9	23			23	1.2	2.5			2.5	9	13	2.1	1.9	52	66	59	64		
24	X	6–5	35	7.			17	13			13	1.8	1.8			1.8	17	17	1.8	1.8	60	56	58	46		
25	Y	7–2	30	19			11	25			25	1.3	2.7			2.7	6	6	2.7	2.1	45	59	52	64		
26	Z	6–10	29	5			9	11			11	1.2	2.1			2.1	14	18	1.8	1.7	47	49	48	52		
27	AA	7–5	39	8			22	14			14	2.3	1.9			1.9	16	16	1.9	2.1	53	45	49	42		
28	BB	7–9	25	13			2	19			19	0.8	2.2			2.2	11	14	2.1	1.6	30	47	39	67		
29	CC	7–2	46	28			28	33			33	3.5	3.6			3.6	2	4	2.9	3.4	62	67	65	55		
30	DD	7–4	51	31			32	36			36	4.3	3.9			3.9	1	2	3.6	4.0	64	68	66	54		

FIG. 93. Classification-data sheet for first grade (inclusion of effort scores)

STEP 14. *Determining the statistical classification*

Before a statistical classification of the pupils in a school can be made, it is necessary that the conditions prevailing in the school shall be determined and provided for in the classification. If the tests are given in one part of the school year the statistical classification will differ from that derived from tests given in another part of the year. Another consideration is the pupils' average rate of learning. If the pupils in a given school, regardless of the length of the school year or the number of years of the elementary-school course, proceed more or less rapidly than normal, account should be taken of the fact. Again, if the reclassification is to be calculated at one time and the reorganization or promotion period at a later time, provisions should be made for that fact.

The technic for taking account of these facts in statistical reclassification is partly contained in a series of tables published by Dr. McCall, who originated them, and is partly derived from the results of the classification data as determined up to this point in Figs. 91 and 92. The steps to be taken in making the classification are as follows:

1. *Determining the mean of the grade scores for promotion separately in all the grades.* The purpose of this step is to make the preliminary investigation necessary to find out the pupils' average rate of learning. If their average rate is greater than normal, the classification should be made upon a different basis than if it is normal or less than normal, as will be shown later. The elementary grades from the first through the sixth, of which the first and fifth have been used for illustration, had the mean G_p scores as given in the first horizontal column of Table LXXVII.

2. *Determining the algebraic differences of the mean G_p scores with their respective norm G_p scores.* As will be shown on page 345, the norm G_p scores are simply the G_p scores of the

grades in the particular month of school in which the tests were given. In this case the tests were given in May of the school year, and the norm G_p scores are shown in the second horizontal line in Table LXXVII.

The algebraic differences are found by merely subtracting the norm G_p scores from the mean G_p scores for each grade, regarding signs. If the norm G_p is less than the mean G_p, the result will be minus; if it is greater, the result will be plus. The differences for these grades are found in the third horizontal line of Table LXXVII.

TABLE LXXVII. DETERMINATION OF MEAN G_p DIFFERENCE

Grades	I	II	III	IV	V	VI
Mean G_p	2.0	2.8	4.1	5.3	6.3	7.3
Norm G_p	1.8	2.8	3.8	4.8	5.8	6.8
Differences	+0.2	0.0	+0.3	+0.5	+0.5	+0.5
Total of differences, +2.0 Mean difference, +0.33						

3. *Determining the mean difference for the G_p scores in all grades.* The mean difference for the G_p scores for all grades is found from the formula

$$\frac{(\text{Sum of } D+) + (\text{sum of } D-)}{\text{Number of differences}} = MnD.$$

The sum of the plus differences added to the sum of the minus differences (regarding signs) and divided by the total number of differences (or grades) is the mean difference desired. In the case here considered the sum of the plus differences is + 2.0, and the sum of the minus differences is 0. Thus the substitution becomes

$$\frac{2.0 + 0}{6} = 0.33.$$

4. *Selecting the proper statistical-classification table.* In the pamphlet of statistical tables for the T-G-B-F-scale system, which every person who plans to work with the T scale should have, there is included a set of three tables for statistical classification.[1] These tables are numbered 2, 3, and 4 in the pamphlet. Table 2 is designed for schools which attempt to cover 0.9 of a standard grade per year; Table 3, for schools which attempt to cover 1.0 of a standard grade per year; and Table 4, for schools which attempt to do 1.1 standard grades of work per year.

The determination of which of these three tables to use for the reclassification of the pupils in any school is decided by the size of the mean G_p difference as found in the third step above. If the difference is less than -0.5, Table 2 should be used; if the difference is from -0.5 to $+0.5$, Table 3 should be used; if the difference is above $+0.5$, Table 4 should be used.[2] In the case shown above the mean difference is $+0.33$, which indicates that Table 3 is applicable here.

The tables are arranged for classification by half-grades or full grades and consist of a series of seven vertical columns. The first column gives the grade score, which, in the case of the classification that has been described, is given as the G_p. Horizontally opposite each grade score on the table are six different classification items, according to the number of months a class has been in the grade at the time that the classification tests are given. Table 3 of McCall's "tables for the T-G-B-F- Scale System" is shown in Table LXXVIII to indicate how it may be used.

[1] The pamphlet *Tables for the T-G-B-F-Scale System* is by Dr. William A. McCall and his students, and is published by Teachers College Bureau of Publications, Columbia University, New York City, from whom it can be purchased at nominal cost.

[2] Compare W. A. McCall and H. H. Bixler's *How to Classify Pupils*, pp. 33–35. Teachers College Bureau of Publications, New York, 1928.

TABLE LXXVIII. TABLE 3 OF MCCALL'S *TABLES FOR THE T-G-B-F-SCALE SYSTEM*

TABLE 3.[1] TABLE FOR THE AUTOMATIC CLASSIFICATION OF PUPILS INTO GRADES OR HALF GRADES ON THE BASIS OF ANY G (GRADE) SCORE. ONE STANDARD GRADE PER YEAR

NUMBER OF MONTHS CLASS HAS BEEN IN HALF OF GRADE IT IS NOW IN							NUMBER OF MONTHS CLASS HAS BEEN IN HALF OF GRADE IT IS NOW IN						
G	0	1	2	3	4	5	G	0	1	2	3	4	5
0.0	KL	KL	KL	KL	KL	KL	4.7	4H	4H	4H	4H	4H	4L
0.1	KL	KL	KL	KL	KL	KL	4.8	5L	4H	4H	4H	4H	4H
0.2	KL	KL	KL	KL	KL	KL	4.9	5L	5L	4H	4H	4H	4H
0.3	KH	KL	KL	KL	KL	KL	5.0	5L	5L	5L	4H	4H	4H
0.4	KH	KH	KL	KL	KL	KL	5.1	5L	5L	5L	5L	4H	4H
0.5	KH	KH	KH	KL	KL	KL	5.2	5L	5L	5L	5L	5L	4H
0.6	KH	KH	KH	KH	KL	KL	5.3	5H	5L	5L	5L	5L	5L
0.7	KH	KH	KH	KH	KH	KL	5.4	5H	5H	5L	5L	5L	5L
0.8	1L	KH	KH	KH	KH	KH	5.5	5H	5H	5H	5L	5L	5L
0.9	1L	1L	KH	KH	KH	KH	5.6	5H	5H	5H	5H	5L	5L
1.0	1L	1L	1L	KH	KH	KH	5.7	5H	5H	5H	5H	5H	5L
1.1	1L	1L	1L	1L	KH	KH	5.8	6L	5H	5H	5H	5H	5H
1.2	1L	1L	1L	1L	1L	KH	5.9	6L	6L	5H	5H	5H	5H
1.3	1H	1L	1L	1L	1L	1L	6.0	6L	6L	6L	5H	5H	5H
1.4	1H	1H	1L	1L	1L	1L	6.1	6L	6L	6L	6L	5H	5H
1.5	1H	1H	1H	1L	1L	1L	6.2	6L	6L	6L	6L	6L	5H
1.6	1H	1H	1H	1H	1L	1L	6.3	6H	6L	6L	6L	6L	6L
1.7	1H	1H	1H	1H	1H	1L	6.4	6H	6H	6L	6L	6L	6L
1.8	2L	1H	1H	1H	1H	1H	6.5	6H	6H	6H	6L	6L	6L
1.9	2L	2L	1H	1H	1H	1H	6.6	6H	6H	6H	6H	6L	6L
2.0	2L	2L	2L	1H	1H	1H	6.7	6H	6H	6H	6H	6H	6L
2.1	2L	2L	2L	2L	1H	1H	6.8	7L	6H	6H	6H	6H	6H
2.2	2L	2L	2L	2L	2L	1H	6.9	7L	7L	6H	6H	6H	6H
2.3	2H	2L	2L	2L	2L	2L	7.0	7L	7L	7L	6H	6H	6H
2.4	2H	2H	2L	2L	2L	2L	7.1	7L	7L	7L	7L	6H	6H
2.5	2H	2H	2H	2L	2L	2L	7.2	7L	7L	7L	7L	7L	6H
2.6	2H	2H	2H	2H	2L	2L	7.3	7H	7L	7L	7L	7L	7L
2.7	2H	2H	2H	2H	2H	2L	7.4	7H	7H	7L	7L	7L	7L
2.8	3L	2H	2H	2H	2H	2H	7.5	7H	7H	7H	7L	7L	7L
2.9	3L	3L	2H	2H	2H	2H	7.6	7H	7H	7H	7H	7L	7L
3.0	3L	3L	3L	2H	2H	2H	7.7	7H	7H	7H	7H	7H	7L
3.1	3L	3L	3L	3L	2H	2H	7.8	8L	7H	7H	7H	7H	7H
3.2	3L	3L	3L	3L	3L	2H	7.9	8L	8L	7H	7H	7H	7H
3.3	3H	3L	3L	3L	3L	3L	8.0	8L	8L	8L	7H	7H	7H
3.4	3H	3H	3L	3L	3L	3L	8.1	8L	8L	8L	8L	7H	7H
3.5	3H	3H	3H	3L	3L	3L	8.2	8L	8L	8L	8L	8L	7H
3.6	3H	3H	3H	3H	3L	3L	8.3	8H	8L	8L	8L	8L	8L
3.7	3H	3H	3H	3H	3H	3L	8.4	8H	8H	8L	8L	8L	8L
3.8	4L	3H	3H	3H	3H	3H	8.5	8H	8H	8H	8L	8L	8L
3.9	4L	4L	3H	3H	3H	3H	8.6	8H	8H	8H	8H	8L	8L
4.0	4L	4L	4L	3H	3H	3H	8.7	8H	8H	8H	8H	8H	8L
4.1	4L	4L	4L	4L	3H	3H	8.8	9L	8H	8H	8H	8H	8H
4.2	4L	4L	4L	4L	4L	3H	8.9	9L	9L	8H	8H	8H	8H
4.3	4H	4L	4L	4L	4L	4L	9.0	9L	9L	9L	8H	8H	8H
4.4	4H	4H	4L	4L	4L	4L	9.1	9L	9L	9L	9L	8H	8H
4.5	4H	4H	4H	4L	4L	4L	9.2	9L	9L	9L	9L	9L	8H
4.6	4H	4H	4H	4H	4L	4L							

[1] Table prepared with the assistance of Grace Moffatt.

Let it be supposed that a pupil has a G_p score of 5.4 and that the tests were given to him in May. He has been in the half of the grade he is now in for three months. In Table 3 of McCall's "Tables for the T-G-B-F-Scale System"[1] it will be seen that the classification given is 5 L, which indicates the first half of the fifth grade as a statistical classification for the pupil. If the tests had been given in March, and the pupil had been in this half of the grade only one month and was to be reclassified immediately, his classification would be in 5 H, or the upper half of the fifth grade. This is found by determining the statistical classification in the column headed 1, indicating one month, directly to the right of G score 5.4.

When the statistical classification has been found from these tables, it should be entered on the classification-data sheet in column 26, which has been reserved for that use. Fig. 94 shows this statistical classification in terms of the G_p score for the fifth grade, and Fig. 95 shows the same data for the first grade.

Pupil 1 has a G_p of 4.7. The tests were given in May, which means that the pupils of this class had been in the half of the grade they are now in for three months at the time the tests were given. The school is also one which attempts to do one standard grade of work a year. Table 3 of McCall's system is used, therefore, and the column labeled "3" in the table is the classification column. For Pupil 1 this column shows for a G_p of 4.7 a statistical classification of 4 H, or the upper half of the fourth grade. Pupil 2, whose G_p is 6.0, is classified as 5 H, or the upper half of the fifth grade. The other G_p scores for all the pupils have been determined in the same way and entered as shown in Figs. 94 and 95 for the fifth and first grades under consideration.

When the reclassification of pupils is to be used as a basis for promotion at some future time, allowance should be made

[1] See Table LXXVIII.

for their development during that period. Thus if a pupil is statistically classified on the first of March in a certain division of a certain grade, and promotion or change is not to come about until June or September, he has four school months of growth before promotion can take place and after the statistical determination of his status has been made according to the means given above.

There is at present no simple method of allowing for such development which is entirely accurate, since the development of individual pupils is dependent upon varying rates of growth. In the majority of cases, however, small injustice will result if the following method is used.

To the G_p score of every pupil add the number of months of growth remaining from the time of the tests to the time of the proposed change or promotion period, and determine a new statistical classification of the pupils from the table previously used. For example, if a pupil were classified in March from a G_p score of 5.8 as 5 H, and promotion were to take place in June, then four months added to his G_p score would mean a classification *for promotion after four months* on the basis of a G_p score of 6.2, which, according to Table LXXVIII, column 3, is 6 L. If this process is followed for a group of pupils, an extra column should be added to the reclassification-data sheet.

STEP 15. *The actual classification*

In making the actual classification of a pupil — that is, in making the final determination of the class, or grade, to which a pupil should be promoted — several principles should be kept in mind. Although it has been found that as a general rule pupils who by reason of their ability are entitled to skip grades seem to fill in the gaps in knowledge with remarkable rapidity, there is a general feeling that too abrupt a progression from a lower grade to a higher is unwise. It will do no

harm for a child to be placed in a higher grade than a normal promotion would place him, but it is unwise to advance him too many grades at once. As a general rule, then, it is unwise to promote a pupil more than one-half grade at a time beyond his actual grade in a school system based on half-yearly promotions. If the pupil is statistically classified several grades beyond his present grade, he should be skipped one grade immediately, allowed to remain there a year or half a year, according to the system of promotions in effect, and then skipped a second time. This system means that the full effects of a reclassification system such as has been described would not be felt for two or three years or possibly longer. The result will be likely to be better for both the pupil and the school.

When a pupil is statistically classified several grades below his present level, it is also unwise to place him immediately in the lower grade where he belongs. It is disheartening to the pupil, discouraging to his parents, and makes trouble for the teacher. If the child is to be demoted, his demotion should be by degrees of one grade or half a grade, according to the plan of promotion in effect in the school system. If he is statistically classified below that point, it is possible to arrange a series of nonpromotions extending over a period of two or more years which, without serious injury to his pride or his effort, will eventually bring him in the grade or class where he really belongs.

A third principle in making the actual classification is to consider the effort previously made by the pupil. If the effort score is less than 50 or greater than 50, indicating that the pupil either has exerted too little effort and is unlikely to be stimulated to his full effort all at once, or that he has expended too much effort and should therefore be restrained, it is probable that his optimum classification is less than his statistical classification.

A fourth principle in making the actual classification is that a thorough study of the pupil himself should be made by both the teacher and the principal before advancing or retarding any pupil beyond what is normal for other pupils. There are many factors — physical, social, and emotional — which might enter to influence the final actual classification and which were neither considered in the statistical classification nor in the data it was made from. The general health of the pupil is one of these considerations. In cases where a pupil has been ill and has not recovered his normal strength or tone, it would probably be unwise to expect of him his full effort, and it would likewise be unwise to place him in as high a grade classification as he might otherwise deserve. Where children have language difficulties which are greater than their lack of knowledge (as is frequently the case with children of foreign parentage), it is probable that they should be classified in a lower grade than their actual achievement might indicate, in order that they may have opportunity to perfect their language facility. There are always cases in schools where emotional difficulties seem to indicate that the pupils should not be placed in grades where the work will leave a gap in knowledge. Some children worry considerably over their inability to understand and have not the temperament to withstand temporary discouragements. Such children should be placed where their work will be an encouragement to them to further effort, even though it is somewhat below their true status.

Teachers are frequently misled in their judgment of the emotional or other characteristics of their pupils. They should consider the statistical classification of pupils seriously before making a decision to retard a pupil in greater proportion than has been indicated in this discussion. The mischievous pupil, or the pupil who seems to lack attentiveness or application, or the pupil whose effort seems to be less

than one would wish, more often than not creates a prejudice against himself in the eyes of the teacher. If the statistical classification shows that such pupils are placed below their normal position in the school, the teacher should suspect that the seeming laziness or inattention or mischievousness is in reality caused by the simplicity of the work of the grade rather than by the pupil's lack of ability to do the work of the higher grade.

Administratively, it is frequently impossible to place all the pupils in their proper grades, or classes. When grades are crowded and there has been a general retardation in the school, pupils can rarely be advanced as quickly as the reclassification would recommend. Under such conditions it would be wise for the younger pupils to be advanced as much as possible, and the older pupils, where the crowding is likely to be greatest, advanced in terms of the room in the various grades. Frequently a school can be relieved of overcrowding if the pupils in the eighth grade who are most advanced can be prepared for high-school entrance in the midyear. By such a process reclassification in a crowded system could be accelerated to take six months rather than a full year.

The recommended actual classifications of the two grades in which the process has been illustrated are shown in column 27 of Figs. 94 and 95, which show the completed classification-data sheets for the fifth and first grades.

Useful data from classification-data sheets. Aside from their use for classifying pupils the data on the classification-data sheets admit of wide use by the classroom teacher, and copies of such sheets should be given to the classroom teacher for such pupils as he may have so that other uses may be made of them. For the classroom teacher there are data available for the individual diagnosis of pupils' difficulties, for the interpretation of teaching, and for determining useful standards for marking. These uses are discussed in the

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
No.	Pu.	CA	M. Sc.	Re. Sc.	Ar. Sc.	Sp. Sc.	T_l	T_r	T_a	T_s	T_e	G_l	G_r	G_a	G_s	G_c	G_c Rk.	Tr. Rk.	G_t	G_p	B_l	B_c	B_p	F_c	St. Cl.	Ac. Cl.
1	A	11–5	72.0	17	16	25	39	42	34	43	40	4.4	4.9	4.7	4.7	4.8	36	33	5.0	4.7	43	44	44	51	4H	5L
2	B	11–9	106.3	20	21	33	49	47	41	49	45	6.4	5.8	5.8	6.2	5.9	21	24	5.6	6.0	52	48	50	46	5H	5H
3	C	10–5	119.3	26	26	41	53	60	50	57	56	7.2	8.8	7.2	8.0	8.1	3	10	6.9	7.5	62	65	64	53	7L	7L
4	D	12–2	100.3	18	24	29	47	43	46	46	45	6.0	5.1	6.5	5.4	5.6	23	28	5.1	5.7	48	46	47	48	5H	5H
5	E	12–0	66.9	16	24	22	38	40	46	40	42	4.2	4.5	6.5	4.2	5.1	26	32	5.0	4.7	40	44	42	54	4H	5L
6	F	13–3	115.2	18	21	18	52	43	41	37	41	7.0	5.1	5.8	3.5	5.1	27	34	4.9	5.8	49	38	44	39	5H	5H
7	G	10–10	145.5	25	24	37	62	58	46	53	53	9.0	8.0	6.5	7.0	7.3	7	4	7.9	8.1	70	61	66	41	7H	7L
8	H	11–10	124.4	23	33	35	55	53	70	51	58	7.6	6.8	11.6	6.6	8.4	2	14	6.3	7.7	58	61	60	53	7H	7L
9	I	12–2	78.1	20	25	22	41	47	48	40	46	4.8	5.8	6.8	4.2	5.9	22	37	4.6	5.2	42	47	45	55	5L	5L
10	J	14–6	109.8	23	25	27	50	53	48	44	50	6.6	6.8	6.8	5.1	6.5	12	15	6.3	6.5	43	43	43	50	6L	6L
11	K	13–0	127.8	21	22	30	55	49	42	47	46	7.6	6.1	6.0	5.6	6.0	17	19	5.9	6.6	53	44	49	41	6L	6L
12	L	10–5	120.5	23	26	36	53	53	50	52	52	7.2	6.8	7.2	6.8	6.9	10	9	7.0	7.0	62	61	62	49	6H	6H
13	M	11–7	128.7	22	21	36	56	51	41	52	48	7.8	6.5	5.8	6.8	6.3	15	7	7.3	7.1	60	52	56	42	6H	6H
14	N	13–6	77.8	14	16	29	41	38	34	46	38	4.8	4.2	4.7	5.4	4.6	37	36	4.8	4.7	37	34	36	47	4H	5L
15	O	12–0	114.0	16	21	22	52	40	41	40	40	7.0	4.5	5.8	4.2	4.9	34	31	5.0	5.8	54	42	48	38	5H	5H
16	P	12–10	82.8	16	22	23	42	40	42	41	41	5.0	4.5	6.0	4.3	5.0	30	35	4.8	5.0	41	40	41	49	4H	5L
17	Q	11–5	107.4	18	25	33	49	43	48	49	46	6.4	5.1	6.8	6.2	5.9	19	18	6.0	6.1	53	50	52	47	5H	5H
18	R	11–3	117.8	23	30	32	52	53	59	48	54	7.0	6.8	9.3	6.0	7.5	6	5	7.7	7.3	57	59	58	52	7L	7L
19	S	10–9	147.6	28	24	47	63	67	46	67	60	9.2	12.6	6.5	11.2	10.3	1	1	10.3	9.9	71	68	70	47	9H	7L
20	T	11–6	93.0	14	23	27	45	38	44	44	41	5.6	4.2	6.2	5.1	5.0	31	29	5.1	5.3	49	45	47	46	5L	5L
21	U	10–8	124.9	25	26	43	55	58	50	60	56	7.6	8.0	7.2	8.8	7.9	4	3	8.1	7.8	63	64	64	51	7H	7L
22	V	11–10	90.3	12	21	31	44	36	41	48	40	5.4	3.8	5.8	5.8	4.8	35	20	5.9	5.3	47	43	45	46	5L	5L
23	W	11–10	93.4	20	23	29	45	47	44	46	46	5.6	5.8	6.2	5.4	5.9	20	23	5.6	5.7	48	49	49	51	5H	5H
24	X	10–1	133.8	25	28	33	57	58	54	49	55	8.9	8.0	8.0	6.2	7.7	5	2	8.4	8.0	68	66	67	48	7H	7L
25	Y	10–9	114.1	23	28	29	52	53	54	46	52	7.9	6.8	8.0	5.4	7.0	9	13	6.4	6.9	60	60	60	50	6H	6H
26	Z	10–9	141.7	23	21	31	60	53	41	48	48	8.6	6.8	5.8	5.8	6.3	14	8	7.2	7.4	68	56	62	38	7L	6H
27	AA	11–2	79.6	12	21	37	41	36	41	53	41	4.8	3.8	5.8	7.0	5.0	32	26	5.1	4.9	47	47	47	50	4H	5L
28	BB	14–4	95.1	16	21	26	46	40	41	44	41	5.8	4.5	5.8	4.9	5.0	33	12	6.5	5.6	52	47	50	45	5L	5L
29	CC	10–3	154.4	23	22	31	66	53	42	48	49	9.7	6.8	6.0	5.8	6.4	13	11	6.7	7.8	76	59	68	33	7H	7L
30	DD	11–2	100.8	22	29	35	47	51	56	51	53	6.0	6.5	8.6	6.6	7.2	8	6	7.5	6.8	53	59	56	56	6H	6H
31	EE	10–7	96.4	22	14	28	46	51	32	45	44	5.8	6.5	4.2	5.2	5.5	25	30	5.5	5.5	54	52	53	48	5L	5L
32	FF	12–5	87.6	16	27	28	43	40	51	45	45	5.2	4.5	7.5	5.2	5.6	24	17	6.0	5.5	44	46	45	52	5L	5L
33	GG	11–3	75.0	17	26	35	40	42	50	51	46	4.6	4.9	7.2	6.6	6.0	18	22	5.9	5.4	45	51	48	56	5L	5L
34	HH	11–0	130.0	22	27	30	56	51	51	47	50	7.8	6.5	7.5	5.6	6.7	11	16	6.2	7.0	63	57	60	44	6H	6H
35	II	11–11	79.4	16	21	30	41	40	41	47	42	4.8	4.5	5.8	5.6	5.1	28	27	5.1	5.0	43	44	44	51	4H	5L
36	JJ	13–0	84.7	18	19	25	43	43	38	43	41	5.2	5.1	5.3	4.7	5.1	29	21	5.9	5.3	41	39	40	48	5L	5L
37	KK	10–8	91.9	21	24	32	44	49	46	48	48	5.4	6.1	6.5	6.0	6.2	16	25	5.5	5.7	52	56	54	54	5H	5H

FIG. 94. Completed classification-data sheet for fifth grade

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
No.	Pu.	CA	M. Sc.	Re. Sc.	Ar. Sc.	Sp. Sc.	T_i	T_r	T_a	T_s	T_e	G_i	G_r	G_a	G_s	G_e	G_e Rk.	Tr. Rk.	G_t	G_p	B_i	B_e	B_p	F_e	St. Cl.	Ac. Cl.
1	A	6–10	34	6			16	12			12	1.7	1.8			1.8	18	20	1.7	1.7	54	50	52	46	1H	1H
2	B	6–3	41	17			24	23			23	2.5	2.5			2.5	10	11	2.2	2.4	69	68	69	49	2L	2L
3	C	6–10	33	2			15	8			8	1.6	1.6			1.6	27	28	1.6	1.6	53	46	50	43	1H	1H
4	D	6–11	28	13			8	19			19	1.1	2.2			2.2	12	1	3.9	2.1	45	56	51	61	2L	2L
5	E	6–4	17	2			−5	8			8	0.5	1.6			1.6	28	30	1.6	1.2	39	52	46	63	1L	1L
6	F	6–11	47	21			29	27			27	3.6	2.9			2.9	5	10	2.5	3.1	66	64	65	48	3L	2H
7	G	7–3	37	7			20	13			13	2.0	1.8			1.8	19	9	2.5	2.0	53	46	50	43	1H	1H
8	H	8–6	11	3			−12	9			9	0.3	1.6			1.6	29	29	1.6	1.1	10	31	21	71	1L	—*
9	I	6–9	22	5			−2	11			11	0.7	1.7			1.7	22	19	1.8	1.3	37	50	44	63	1L	1L
10	J	6–8	40	11			23	17			17	2.4	2.1			2.1	15	8	2.5	2.3	63	57	60	44	2L	2L
11	K	6–2	38	5			20	11			11	2.2	1.7			1.7	23	23	1.7	1.9	67	58	63	41	1H	1H
12	L	7–0	40	26			23	31			31	2.4	3.4			3.4	3	5	2.9	2.9	59	67	63	58	2H	2H
13	M	6–5	27	4			6	10			10	1.0	1.7			1.7	24	21	1.7	1.4	49	53	51	54	1L	1L
14	N	6–6	34	2			16	8			8	1.7	1.6			1.6	30	25	1.6	1.6	58	50	54	42	1H	1L
15	O	7–7	37	21			20	27			27	2.2	2.9			2.9	4	3	3.4	2.7	50	57	54	57	2H	2H
16	P	6–4	29	17			11	23			23	1.3	2.5			2.5	7	12	2.2	2.0	55	67	61	62	1H	1H
17	Q	7–2	32	11			13	17			17	1.5	2.1			2.1	13	15	2.1	1.9	47	51	49	54	1H	1H
18	R	7–0	28	2			8	8			8	1.1	1.6			1.6	25	26	1.6	1.4	44	44	44	50	1L	1L
19	S	6–10	22	2			−2	8			8	0.7	1.6			1.6	26	22	1.7	1.3	36	46	41	60	1L	1L
20	T	6–7	22	4			−2	10			10	0.7	1.7			1.7	20	24	1.7	1.3	39	51	45	62	1L	1L
21	U	6–0	36	16			19	22			22	1.9	2.5			2.5	8	7	2.5	2.3	69	72	71	53	2L	2L
22	V	6–11	23	4			0	10			10	0.7	1.7			1.7	21	27	1.6	1.3	37	47	42	60	1L	1L
23	W	6–5	29	17			9	23			23	1.2	2.5			2.5	9	13	2.1	1.9	52	66	59	64	1H	1H
24	X	6–5	35	7			17	13			13	1.8	1.8			1.8	17	17	1.8	1.8	60	56	58	46	1H	1H
25	Y	7–2	30	19			11	25			25	1.3	2.7			2.7	6	6	2.7	2.1	45	59	52	64	2L	1H
26	Z	6–10	29	5			9	11			11	1.2	2.1			2.1	14	18	1.8	1.7	47	49	48	52	1H	1H
27	AA	7–5	39	8			22	14			14	2.3	1.9			1.9	16	16	1.9	2.1	53	45	49	42	2L	1H
28	BB	7–9	25	13			2	19			19	0.8	2.2			2.2	11	14	2.1	1.6	30	47	39	67	1H	1L
29	CC	7–2	46	28			28	33			33	3.5	3.6			3.6	2	4	2.9	3.4	62	67	65	55	3L	2H
30	DD	7–4	51	31			32	36			36	4.3	3.9			3.9	1	2	3.6	4.0	64	68	66	54	3H	2H

FIG. 95. Completed classification-data sheet for first grade

*Special school.

chapters on diagnosis and on the use of tests in teaching. In addition, for the use of the classroom teacher the reclassification data is more valuable when supplemented by B and F scores, not alone in composites but also in the separate subjects. Thus a list of B_r, B_a, B_s, and F_r, F_a, and F_s scores might become a valuable aid in teaching.

For the supervisor there is much information with respect to subject accomplishments of the pupils and to general grade averages and grade levels, and for the administrator there is information relating to grade and subject levels. These are discussed in Chapter XVI.

Classification of a school by means of the age scale. Classification of a school can be accomplished satisfactorily by means of the age scale only, if it is so desired. Since the technic of classification has been rather minutely described in terms of the T scale, it is not necessary to follow the process so closely. In Fig. 96 is shown a classification-data sheet for the same class as is shown in Fig. 94, with the classification based upon age units. The results of this type of classification are essentially the same as those on the T scale, but, as will be seen, the data are somewhat less easy to interpret and are somewhat less desirable to use. The process of classification (assuming that the same series of tests was given as in the case of the pupils involved in the data of Fig. 94 and that the same preparations for the classification were made) is as follows, in terms of the columns on the classification-data sheet:

Column 1. Pupils' numbers. When the pupils involved in the classification for any class have been tested, their names should be placed in alphabetical order and numbered, as in Fig. 96, column 1.

Column 2. Pupils' names in alphabetical order. The names of the pupils in the class to be classified should be placed in alphabetical order, as shown in the second column of Fig. 96.

Column 3. Chronological ages. The ages of the pupils calculated as recommended in Chapter V should be placed in column 3, as shown in Fig. 96.

Column 4. Raw scores in the intelligence test. When the intelligence test has been scored and re-checked, the raw scores should be placed opposite the names of the pupils in column 4.

Columns 5–7. Raw scores on educational tests. When the tests which are to be used for the educational aspects of reclassification have been given and scored, the raw scores for the pupils should be placed in the next three columns. Column 4 of Fig. 96 contains the raw scores received by the pupils in the reading test, column 5 the raw scores received in the arithmetic test, and column 6 the raw scores received in the spelling test.

Column 8. Mental age. In column 8 the mental ages of the pupils, as determined by the intelligence test, should be recorded. The mental age is found, as a rule, by consulting the table of mental ages which accompanies the test. The mental ages in column 8 have been determined by locating the raw scores in Table XXXIV, p. 259, for the National Intelligence Test, Form *A*. Where such a table is not available, the age norms for the test should be made into an age table according to the method given in Chapter IX. When this table has been constructed, the raw scores of the pupils in an entire school can easily be transposed into mental ages. It should also be noted that the mental ages can be found if the intelligence quotients and the chronological ages are known. This process should be used when an intelligence test has been given some time before the time of reclassification. The mental age should always be placed on the classification-data sheet corrected to the time of the giving of the educational tests; therefore, if there are mental ages of a previous period available, it is unnecessary to give a new

mental test, but it is necessary to correct the mental ages for the time which has elapsed.[1]

Column 9. Reading age. The reading age for the Thorndike-McCall Reading Scale is found from a table of reading-age norms as given in Table 3 of the manual of directions for that test. Here the raw score is turned into a T score, which in turn is read directly by Table 3 as a reading age. For example, by Table 1 of the Thorndike-McCall manual of directions a raw score of 21 is a T score of 46; by Table 3 of the manual a T score of 46 is equivalent to a reading age of 138 months, or 11–6.

When such a table of age scores is not available for a test, the table can be constructed according to the method given on page 212; and when the table of age scores for the test is completed, it can be used for transmuting the raw scores for the entire school.

Column 10. Arithmetic age. The arithmetic ages found in column 10 have been derived from the age norms for the Woody-McCall Arithmetic Test as given in the sheet of directions. These norms were converted by the method described for transposing age norms into a raw-score-age table.[2]

Column 11. Spelling age. The spelling age was determined by transmuting the raw scores in spelling, found in column 7, into age scores by means of a table of raw-score-spelling ages, constructed as shown on pages 212–218, which used the age norms as given in the Morrison-McCall Spelling Test directions.

Column 12. Educational age. The educational age is a single score which combines the different subject ages which have been determined and are recorded in columns 9–11. This is a weighted-average, determined in the same way as the weighted-average T score in education described on page 318.

[1] See page 263. [2] See page 212.

In this case the formula was

$$\frac{3\ ReA + 2\ ArA + SpA}{6} = EA.$$

For Pupil 1 the substitution becomes

$$\frac{3(132) + 2(130) + 130}{6} = 131.$$

Column 13. Rank of educational ages. In column 13 should be inserted the ranking of the educational ages, in which the highest educational age is given a rank of 1, the next highest a rank of 2, and so on. When there are two or more ages of the same size, it is merely necessary to give them different ranks as they are found, since the ranking is used merely toward helping to determine the data to be found in column 15.

Column 14. Ranking by teacher for promotion. As in the previous form of classification, a teacher should be requested to give a ranking of all the pupils in terms of their worthiness for promotion. Each pupil should have a definite ranking, with the most worthy allotted the rank of 1, the next most worthy the rank of 2, and so on. For this classification the rankings made by the teacher are as given in column 14 of Fig. 96.

Column 15. Age status for promotion by teacher. The age status of worthiness for promotion is determined by transmutation from the educational ages by means of the rankings made by the teacher, in the same way as the grade status by the teacher was determined in the T-scale form for classification. The teacher should allot to the pupil whom he ranks 1 an age status equivalent to the educational age found in column 12, achieved by the pupil ranking 1 in column 13. Table LXXIX shows in column 1 the names of five pupils, in column 2 their educational ages, in column 3 their rankings in educational ages, in column 4 their rankings by the teacher, and in column 5 the transmuted age-status ranking as deter-

mined by the teacher. The arrows show the derivation of the scores in column 5 from the scores in column 2. When a similar procedure is carried out for the scores in column 12 of Fig. 96, the result will be as shown in column 15 of the same figure.

Column 16. Promotion age. Three different age scores have been determined up to this point — the mental age, a general educational age, and an age which shows the judgment of the teacher as to the worthiness of the pupils for promotion. A combination of these three scores, weighted in terms of their values, will give a general score for promotion purposes. This is called a promotion age.

TABLE LXXIX. TRANSMUTATION OF EA TO TrA

Pupils	EA	EA Rank	Tr Rank	TrA
A	151	2	1	160
B	126	5	4	127
C	135	3	2	151
D	160	1	3	135
E	127	4	5	126

Since the least reliable of these three ages is that transmuted from the educational age into an age status as determined by the teacher, and since both mental age and educational age were derived from carefully handled tests, it is probable that these last should be given more weight. A suggested weighting was given in the previous type of classification, which, for this classification, is expressed in the formula

$$\frac{2\,MA + 2\,EA + TrA}{5} = PrA.$$

The substitution for Pupil 1 becomes

$$\frac{2(123) + 2(131) + 134}{5} = 128.$$

When all the promotion ages have been determined they should be entered in column 16, as shown in Fig. 96.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
No.	Pu.	CA	M. Sc.	Re. Sc.	Ar. Sc.	Sp. Sc.	MA	Re. A.	Ar. A.	Sp. A.	EA	EA Rk.	Tr. Rk.	Tr. A.	Pr. A.	Pr. Q.	AR	St. Cl.	Ac. Cl.
1	A	137	72.0	17	16	25	123	132	130	130	131	36	33	134	128	93	106	4L	5L
2	B	141	106.3	20	21	33	147	144	144	148	145	18	24	142	145	103	98	5H	5H
3	C	125	119.3	26	26	41	158	177	161	171	171	2	10	158	163	130	108	7L	6H
4	D	146	100.3	18	24	29	143	135	153	139	142	23	28	135	141	96	99	5L	5L
5	E	144	66.9	16	24	22	118	128	153	124	136	26	32	134	128	88	115	4L	5L
6	F	159	115.2	18	21	18	155	135	144	116	135	28	34	133	143	90	87	5H	5H
7	G	130	145.5	25	24	37	182	171	153	159	163	6	4	169	172	132	89	7H	7L
8	H	142	124.4	23	33	35	162	156	196	154	169	3	14	151	163	114	104	7L	7L
9	I	146	78.1	20	25	22	128	144	156	124	145	19	37	129	135	92	113	5L	5L
10	J	174	109.8	23	25	27	150	156	156	135	153	12	15	150	151	86	102	6L	6L
11	K	156	127.8	21	22	30	166	147	146	142	146	17	19	145	154	98	87	6L	6L
12	L	125	120.5	23	26	36	159	156	161	156	158	9	9	158	158	126	99	6H	6H
13	M	139	128.7	22	21	36	167	153	144	156	151	13	7	162	159	116	90	6H	6H
14	N	162	77.8	14	16	29	127	124	130	139	129	37	36	131	129	80	101	4H	5L
15	O	144	114.0	16	21	22	154	128	144	124	133	34	31	134	142	98	86	5H	5H
16	P	154	82.8	16	22	23	131	128	146	126	134	30	35	132	132	85	102	4H	5L
17	Q	137	107.4	18	25	33	148	135	156	148	144	22	18	145	146	106	97	5H	5H
18	R	135	117.8	23	30	32	156	156	180	146	162	7	5	167	161	111	103	7L	7L
19	S	129	147.6	28	24	47	185	201	161	192	186	1	1	186	186	144	100	8H	7L
20	T	138	93.0	14	23	27	138	124	149	135	134	31	29	135	136	98	90	5L	5L
21	U	128	124.9	25	26	43	162	171	161	177	169	4	3	169	166	129	104	7L	7L
22	V	142	90.3	12	21	31	136	119	144	144	132	35	20	145	136	95	97	5L	5L
23	W	142	93.4	20	23	29	138	144	149	139	145	20	23	142	142	100	105	5H	5H
24	X	121	133.8	25	28	33	169	171	171	148	167	5	2	171	169	139	98	7H	7L
25	Y	129	114.1	23	28	29	154	156	171	139	158	10	13	151	155	120	102	6H	6L
26	Z	129	141.7	23	21	31	177	156	144	144	150	15	8	161	163	126	84	7L	6H
27	AA	134	79.6	12	21	37	128	119	144	159	134	32	26	136	132	98	104	4H	5L
28	BB	133	95.1	16	21	26	139	128	144	132	134	33	12	153	140	105	96	5L	5L
29	CC	123	154.4	23	22	31	191	156	146	144	151	14	11	155	168	136	79	7H	7L
30	DD	134	100.8	22	29	35	143	153	176	154	161	8	6	163	154	115	112	6L	6L
31	EE	127	96.4	22	14	28	140	153	124	137	141	25	30	134	139	115	100	5L	5L
32	FF	149	87.6	16	27	28	134	128	165	137	142	24	17	146	140	94	106	5L	5L
33	GG	135	75.0	17	26	35	126	132	161	154	145	21	22	144	137	101	115	5L	5L
34	HH	132	130.0	22	27	30	168	153	165	142	155	11	16	149	159	128	92	6H	6H
35	II	143	79.4	16	21	30	128	128	144	142	136	27	27	136	133	93	106	4H	5L
36	JJ	156	84.7	18	19	25	132	135	138	130	135	29	21	145	136	87	102	5L	5L
37	KK	128	91.9	21	24	32	137	147	153	146	149	16	25	141	143	111	108	5H	5H

FIG. 96. Completed age-scale classification-data sheet for fifth grade

Column 17. Promotion quotient. The promotion quotient is a score which shows the relationship of the work and status of a pupil as compared with pupils of his own chronological age. It is a good score for determining the probable rate of growth that can be expected of a pupil. If the quotient is above 100, it indicates that the pupil, regardless of his general age status, has moved more rapidly than normal pupils of his own age; if it is less than 100, it indicates that he has progressed less rapidly than normal pupils of his own age.

The promotion quotient is determined by dividing the promotion age of a pupil by his chronological age. The CA is found in column 3 and the PrA is found in column 16 of Fig. 96. Decimal points are eliminated for ease in reading. The formula is

$$(100 \times) \frac{PrA \text{ (column 16)}}{CA \text{ (column 7)}} = PrQ \text{ (column 17)}.$$

The substitution for Pupil 1 is as follows:

$$(100 \times) \frac{128}{137} = 93.$$

When all the promotion quotients have been determined they should be included in column 17, as shown in Fig. 96.

Column 18. Achievement ratio. The achievement ratio, as described in Chapter XI, is a score which shows the relationship of a pupil's achievement to what he might have done had his work been normal; it is therefore a measure of individual effort. The formula for the achievement ratio, where the ratio is multiplied by 100 to eliminate decimal points, is

$$(100 \times) \frac{EA \text{ (column 12)}}{MA \text{ (column 8)}} = AR \text{ (column 18)}.$$

The calculation for Pupil 1 is

$$(100 \times) \frac{131}{123} = 106.$$

Column 19. The statistical classification. The statistical classification should be made upon the basis of the promotion ages as found in column 16. The basis for transmuting these ages into grade units is an age-grade transmutation table for classification. Any age-grade transmutation table which best fits a particular school system can be used, but Table LXXX, which is based upon general conditions for schools attempting to do one standard grade of work a year, has more universal value.

This age-grade classification table is based upon a base age at the beginning of the first grade of eighty months and the thirteen-month period as contained in Table XXVIII A and Table XXVIII B, and is similar to that used by Dr. McCall in constructing his Table 3.[1] In school systems which have previously been reclassified this table will undoubtedly be found to be too high, and one better suited to the local conditions should be constructed which is based on a more desirable basis from Table XXVIII.

The table is constructed for average schools where normal retardation is found, and can be used for classification either by grades or by half-grades. The columns of the table are headed by figures indicating the number of months the pupils have been in the half-grade they are now in. To use the table, the promotion ages should be found in the proper column, and the number opposite in the first column will show the proper classification.

For the classification of the pupils in Fig. 96, column 3 (since the pupils had been in the grade three months) of Table LXXX was used. By this table Pupil 1, with a PrA of 128, is shown to be statistically classified in grade 4 L; Pupil 2, with a PrA of 145, in grade 5 H; and so on, as shown in column 19, Fig. 96.

[1] See Table LXXVIII, p. 344.

TABLE LXXX. AGE-GRADE RECLASSIFICATION TABLE (BASED UPON GRADES AND HALF-GRADES FOR AVERAGE SCHOOLS)

GRADES	MONTHS IN HALF-GRADE					
	0	1	2	3	4	5
I L	80–86	81–87	82–88	83–89	84–90	85–91
I H	87–92	88–93	89–94	90–95	91–96	92–97
II L	93–99	94–100	95–101	96–102	97–103	98–104
II H	100–105	101–106	102–107	103–108	104–109	105–110
III L	106–112	107–113	108–114	109–115	110–116	111–117
III H	113–118	114–119	115–120	116–121	117–122	118–123
IV L	119–125	120–126	121–127	122–128	123–129	124–130
IV H	126–131	127–132	128–133	129–134	130–135	131–136
V L	132–138	133–139	134–140	135–141	136–142	137–143
V H	139–144	140–145	141–146	142–147	143–148	144–149
VI L	145–151	146–152	147–153	148–154	149–155	150–156
VI H	152–157	153–158	154–159	155–160	156–161	157–162
VII L	158–164	159–165	160–166	161–167	162–168	163–169
VII H	165–170	166–171	167–172	168–173	169–174	170–175
VIII L	171–177	172–178	173–179	174–180	175–181	176–182
VIII H	178–183	179–184	180–185	181–186	182–187	183–188

Column 20. Actual classification. The actual classification of pupils by means of the age scale should take into account the principles of classification already outlined. Pupils should be retarded or accelerated according to the statistical classification in terms of a promotion or a demotion of not more than one grade when the system has half-yearly promotions. The classification of pupils into classes can be accomplished according to the statistical classification within these limits or else according to the promotion quotient. This classification takes into account the rate at which pupils progress as well as their general level for promotion. When there are two classes of the same grade, such as 4 L and 4 L, both of which are beginning at the same point, the distribution of pupils in the two classes should be made on the basis of the promotion quotient. The pupils should be listed in terms of their promotion quotients (highest promotion quo-

tients first), and one class can be made up of the higher half of the quotients for the grade, and the other can be made up of the lower half. This will give two groups, each containing pupils classified in terms of achievement as well as probable rate of learning.

In making final assignments to grades the achievement ratio should be consulted. This ratio indicates the degree of effort previously expended by the pupil in terms of his ability, and is a convenient measure of his probable future effort. If the AR is high, a pupil is probably classified statistically too high, and his actual classification should be lower; if the AR is low, it is probable that the statistical classification is higher than the optimum placement of the pupil, and that the actual classification should therefore be lower. It is only when the AR is approximately 100, and it is thereby understood that the pupil has been moving normally for one of his ability, that the statistical classification is likely to place him in his optimum position in the school.

Classification of a school by means of both age and grade scales. It is somewhat more convenient and also reduces somewhat the tediousness of calculation to classify a school by means of both grade norms and ages. To do this the classification-data sheet used and shown as Fig. 96 must be slightly modified by the addition of the grade levels of the pupils. When these levels are used, the average scores used for promotion can be found in terms of grade scores rather than age scores, and the statistical classification can be made by means of Dr. McCall's Statistical Classification Tables, as used when classifying by means of the T-scale system.

Fig. 97 shows a classification-data sheet which includes both age and grade scores for the pupils. The process of classification can be described briefly in terms of the columns of data included on this sheet.

Column 1. Pupil's number. This is found as in the other forms of classification, p. 313.

Column 2. Pupil's name. The names of the pupils are listed in alphabetical order. The fact that letters are used in the illustration instead of names should cause no confusion.

Column 3. Chronological age. The chronological ages of the pupils are calculated according to the directions given on page 95, and are placed in this column.

Columns 4–7. Raw scores from tests. In column 4 should be placed the raw scores from the intelligence, or ability, test, in column 5 the raw scores from the reading test, in column 6 the raw scores from the arithmetic test, and in column 7 the raw scores from the spelling test.

Columns 8–11. Age scores. In column 8 should be listed the mental ages of the pupils as determined on page 353; in column 9 should be listed the reading ages as transmuted from the raw scores in reading, as described on page 212; in column 10 should be placed the arithmetic ages as derived from the raw scores by the method given on page 212; in column 11 should be listed the spelling ages as derived according to the method given on page 212.

Column 12. Educational age. This is a weighted average age, according to the best judgment of the examiner. The weighting used in Fig. 97 is the same as that used for the similar age on Fig. 96. The formula was

$$\frac{3\,(ReA) + 2\,ArA + SpA}{6} = EA.$$

Column 13. Grade status in intelligence. This score is found from the age-grade table by the same method that was used on page 319, when classifying by the T scale. It can be determined either from the raw scores of column 4 or from the mental ages of column 8, depending upon which transmutation table is more easily available.

Column 14. Grade score in reading. This is determined according to the method described on page 203. It can be derived either from column 5, which contains the raw scores in reading, or from column 9, which contains the reading ages. The direction manual for the test which was here used gives tables for either transmutation, and it will probably be found convenient to record both age and grade scores at the same time directly from the raw scores.

Column 15. Grade scores in arithmetic. This score is found as in the T-scale classification on page 319. It can be found either through the raw scores of column 6 or from the arithmetic ages in column 10.

Column 16. Grade score in spelling. This score is found as in the T-scale classification given on page 319, and may be determined either from the raw scores in spelling in column 7 or from the spelling-age scores in column 11.

Column 17. Grade score in education. This is a composite weighted score derived from the three preceding subject-matter grade scores. The weighting may be determined according to the desires of the teacher or examiner. The weighting used here has been that recommended on page 319 for the T-scale classification, in which the following formula was used:

$$\frac{3\,G_r + 2\,G_a + G_s}{6} = G_e.$$

The substitutions are made in the way previously described.

Column 18. Rank of grade score in education. This is the ranking from the best to the least of the various grade scores contained in column 18. The purpose for the ranking is the same as that used to transpose the ratings of the teacher into a grade status comparable with the other grade scores previously obtained.

Column 19. Rank by teacher of pupils in terms of worthiness for promotion. The teacher is asked to rank the pupils from

best to poorest in terms of worthiness for promotion, and the ranks are then placed in column 19, as shown in Fig. 97.

Column 20. Grade status according to teacher's ranking. The grade status according to the teacher's ranking for promotion is found as described on page 324 by giving to the pupil with the Tr.Rk. of 1 the G_t equivalent to the G_e of the pupil with a G_eRk. of 1, to the pupil with a Tr.Rk. of 2 the G_t equivalent to the G_e of the pupil with a G_eRk. of 2, and so on.

Column 21. Grade status for promotion. This score is a weighted composite score of the grade status in intelligence (G_i, column 13), the grade status in education (G_e, column 17), and the grade status according to the teacher's ranking (G_t, column 20). The weighting here used gives greater emphasis to the two measures resulting from testing than it gives to the measure resulting from the opinion of the teacher. It has been described on page 326, and the formula is

$$\frac{2\,G_i + 2\,G_e + G_t}{5} = G_p.$$

Column 22. Promotion age. The promotion age, which is found in order to calculate the promotion quotient, is determined by adding the mental and educational ages and dividing by 2. These ages are found in column 22.

Column 23. Promotion quotient. In the method for classification by the age scale it was possible to find a promotion quotient which included the various ages used in determining the promotion age.[1] That is not possible in this classification scheme, since the ranking by the teacher has been transposed into a grade score rather than an age score. The promotion quotient, however, is a valuable asset in determining the class status of an individual within a grade, and is therefore worth finding. It can be found from

[1] See page 358.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
No.	Pu.	CA	M. Sc.	Re. Sc.	Ar. Sc.	Sp. Sc.	MA	Re. A.	Ar. A.	Sp. A.	EA	G_i	G_r	G_a	G_s	G_e	G_e Rk.	Tr. Rk.	G_t	G_p	Pr. A.	Pr. Q	AR	St. Cl.	Ac. Cl.
1	A	137	72.0	17	16	25	123	132	130	130	131	4.4	4.9	4.7	4.7	4.8	36	33	5.0	4.7	127	92	106	4H	5L
2	B	141	106.3	20	21	33	147	144	144	148	144	6.4	5.8	5.8	6.2	5.9	21	24	5.6	6.0	146	103	98	5H	5H
3	C	125	119.3	26	26	41	158	177	161	171	171	7.2	8.8	7.2	8.0	8.1	3	10	6.9	7.5	165	132	108	7L	6L
4	D	146	100.3	18	24	29	143	135	153	139	142	6.0	5.1	6.5	5.4	5.6	23	28	5.1	5.7	143	97	99	5H	5H
5	E	144	66.9	16	24	22	118	128	153	124	136	4.2	4.5	6.5	4.2	5.1	26	32	5.0	4.7	127	88	115	4H	5L
6	F	159	115.2	18	21	18	155	135	144	116	139	7.0	5.1	5.8	3.5	5.1	27	34	4.9	5.8	147	92	87	5H	5H
7	G	130	145.5	25	24	37	182	171	153	159	163	9.0	8.0	6.5	7.0	7.3	7	4	7.9	8.1	173	133	89	7H	7L
8	H	142	124.4	23	33	35	162	156	196	154	169	7.6	6.8	11.6	6.6	8.4	2	14	6.3	7.7	166	116	104	7H	7L
9	I	146	78.1	20	25	22	128	144	156	124	152	4.8	5.8	6.8	4.2	5.9	22	37	4.6	5.2	140	96	113	5L	5L
10	J	174	109.8	23	25	27	150	156	156	135	153	6.6	6.8	6.8	5.1	6.5	12	15	6.3	6.5	152	87	102	6L	6L
11	K	156	127.8	21	22	30	166	147	146	142	146	7.6	6.1	6.0	5.6	6.0	17	19	5.9	6.6	156	100	87	6L	6L
12	L	125	120.5	23	26	36	159	156	161	156	158	7.2	6.8	7.2	6.8	6.9	10	9	7.0	7.0	159	119	99	6H	6H
13	M	139	128.7	22	21	36	167	153	144	156	151	7.8	6.5	5.8	6.8	6.3	15	7	7.3	7.1	159	114	90	6H	6H
14	N	162	77.8	14	16	29	127	124	130	139	129	4.8	4.2	4.7	5.4	4.6	37	36	4.8	4.7	128	79	101	4H	5L
15	O	144	114.0	16	21	22	154	128	144	124	133	7.0	4.5	5.8	4.2	4.9	34	31	5.0	5.8	144	100	86	5H	5H
16	P	154	82.8	16	22	23	131	128	146	126	133	5.0	4.5	6.0	4.3	5.0	30	35	4.8	5.0	132	86	102	4H	5L
17	Q	137	107.4	18	25	33	148	135	156	148	144	6.4	5.1	6.8	6.2	5.9	19	18	6.0	6.1	146	106	97	5H	5H
18	R	135	117.8	23	30	32	156	156	180	146	162	7.0	6.8	9.3	6.0	7.5	6	5	7.7	7.3	159	117	103	7L	7L
19	S	129	147.6	28	24	47	185	201	161	192	186	9.2	12.6	6.5	11.2	10.3	1	1	10.3	9.9	186	144	100	9H	7L
20	T	138	93.0	14	23	27	138	124	149	135	134	5.6	4.2	6.2	5.1	5.0	31	29	5.1	5.3	136	98	90	5L	5L
21	U	128	124.9	25	26	43	162	171	161	177	169	7.6	8.0	7.2	8.8	7.9	4	3	8.1	7.8	166	129	104	7H	7L
22	V	142	90.3	12	21	31	136	119	144	144	132	5.4	3.8	5.8	5.8	4.8	35	20	5.9	5.3	134	94	97	5L	5L
23	W	142	93.4	20	23	29	138	144	149	139	145	5.6	5.8	6.2	5.4	5.9	20	23	5.6	5.7	142	100	105	5H	5H
24	X	121	133.8	25	28	33	169	171	171	148	167	8.0	8.0	8.0	6.2	7.7	5	2	8.4	8.0	168	138	98	7H	7L
25	Y	129	114.1	23	28	29	154	156	171	139	158	7.0	6.8	8.0	5.4	7.0	9	13	6.4	6.9	156	120	102	6H	6H
26	Z	129	141.7	23	21	31	177	156	144	144	150	8.6	6.8	5.8	5.8	6.3	14	8	7.2	7.4	164	127	84	7L	6H
27	AA	134	79.6	12	21	37	128	119	144	159	134	4.8	3.8	5.8	7.0	5.0	32	26	5.1	4.9	131	97	104	4H	5L
28	BB	133	95.1	16	21	36	139	128	144	132	134	5.8	4.5	5.8	4.9	5.0	33	12	6.5	5.6	137	103	96	5L	5L
29	CC	123	154.4	23	22	31	191	156	146	144	151	9.7	6.8	6.0	5.8	6.4	13	11	6.7	7.8	171	139	79	7H	7L
30	DD	134	100.8	22	29	35	143	153	176	154	161	6.0	6.5	8.6	6.6	7.2	8	6	7.5	6.8	152	113	112	6H	6H
31	EE	127	96.4	22	14	28	140	153	124	137	141	5.8	6.5	4.2	5.2	5.5	25	30	5.0	5.5	141	111	100	5L	5L
32	FF	149	87.6	16	27	28	134	128	165	137	142	5.2	4.5	7.5	5.2	5.6	24	17	6.0	5.5	138	92	106	5L	5L
33	GG	135	75.0	17	26	35	126	132	161	154	145	4.6	4.9	7.2	6.6	6.0	18	22	5.9	5.4	136	100	115	5L	5L
34	HH	132	130.0	22	27	30	168	153	165	142	155	7.8	6.5	7.5	5.6	6.7	11	16	6.2	7.0	162	122	92	6H	6H
35	II	143	79.4	16	21	30	128	128	144	142	136	4.8	4.5	5.8	5.6	5.1	28	27	5.1	5.0	132	92	106	4H	5L
36	JJ	156	84.7	18	19	25	132	135	138	130	135	5.2	5.1	5.3	4.7	5.1	29	21	5.9	5.3	134	85	102	5L	5L
37	KK	128	91.9	21	24	32	137	147	153	146	149	5.4	6.1	6.5	6.0	6.2	16	25	5.5	5.7	143	111	108	5H	5H

FIG. 97. Completed age-grade scale classification-data sheet for fifth grade

the composite of the mental and educational ages according to the formula

$$(100 \times) \frac{PrA}{CA} = PrQ.$$

When the promotion quotients are determined they can be placed in column 23.

Column 24. Achievement ratio. This is the score which indicates the quality of effort for a pupil, and should be used in making the actual classification. It is found according to the formula

$$(100 \times) \frac{EA}{MA} = AR.$$

Column 25. Statistical classification. This is a classification of a pupil according to his grade status for promotion (column 21). It is determined by means of the McCall classification table. When the statistical classification has been completed, the grades can be listed in column 25.

Column 26. Actual classification. The actual classification is a modification of the statistical classification according to the method described on page 346 for the T-scale classification.

Chapter summary. The classification of a school involves the problem of placing pupils in grades and classes where they can do their best work. This usually means placing those of equal achievement, mentality, and rates of growth in the same divisions or classes.

The process of classification is a procedure whereby the results of testing are converted into scores which can be combined, weighted, and treated statistically in order to determine achievement, mentality, and rate.

Classification can be accomplished through using the T scale, the age scale, or the age and grade scales combined. In the T-scale classification the raw scores on the tests are determined and are transposed into T scores, and the T scores are then transmuted into grade scores. A weighting of the

T scores is made to determine a weighted composite T score in education, and a similar weighting is made to find a general grade-score average. A technic is given for transposing the ranking of a teacher into a grade score comparable with the other grade scores found. The weighted grade scores are then combined to find a single grade score for promotion, which in turn is used to determine, by means of a classification table, a statistical classification. To facilitate the modification of the statistical classification, other scores, such as the ability and effort scores, are also found. The entire process results in the placement of pupils in their optimum grade or class classification.

The process outlined for the T scale is possible with the age scale or the age-grade scale, with certain slight modifications in the use of the scores, but with quite similar results.

SELECTED BIBLIOGRAPHY

BUCKINGHAM, B. R. Research for Teachers, pp. 258–294. Silver, Burdett and Company, New York, 1926.

GREGORY, C. A. Fundamentals of Educational Measurement, pp. 190–192. D. Appleton and Company, New York, 1922.

MCCALL, W. A., and BIXLER, H. H. How to Classify Pupils. Teachers College Bureau of Publications, New York, 1928.

MCCALL, W. A. How to Measure in Education, pp. 19–66. The Macmillan Company, New York, 1922.

MONROE, W. S., DEVOSS, J. C., and KELLY, F. J. Educational Tests and Measurements (Revised and Enlarged Edition), p. 437. Houghton Mifflin Company, Boston, 1924.

PAULU, E. M. Diagnostic Testing and Remedial Teaching, pp. 55–67. D. C. Heath & Co., Boston, 1924.

TERMAN, L. M. The Intelligence of School Children, pp. 72–74. Houghton Mifflin Company, Boston, 1919.

TRABUE, M. R. Measuring Results in Education, p. 444. American Book Company, New York, 1924.

WILSON, G. M., and HOKE, K. J. How to Measure (Revised Edition), pp. 356–358, 378–384. The Macmillan Company, New York, 1928.

CHAPTER XIV

THE DIAGNOSIS OF INDIVIDUALS

Types of educational diagnosis. Educational diagnosis is one of the most complicated and extensive processes in the field of education. It is intimately connected with every process of learning, with every portion of the educational environment, and with every activity of the teacher. It is concerned with the development of learning, with the acceleration of learning, with the correction of learning, and with the guidance of learning. It can be concerned with every pupil in school, with every class, and with every school system. As such, diagnosis must be selective. It must be concerned with those elements where it is most greatly needed and where it will bring the greatest results. It must be casual on occasion; it must be intensive when necessary. It is needed with all individuals at some time during their educational development, and is needed with some individuals almost constantly.

Educational diagnosis may be concerned with an individual pupil. As such, it is largely a relationship between the teacher and the pupil. The teacher uses the technics of individual diagnosis to help in the adjustment of the pupil, to find his difficulties, and to locate the defects in his learnings. Individual educational diagnosis is a process of individual adjustment.

Educational diagnosis may also be concerned with the classroom. When it is so concerned, it consists in a relationship between a supervisor or principal and the classroom teacher. Here the technics of educational diagnosis are used

to aid the teacher in the adjustment of his work for the improvement of class instruction, for the maintenance of minimal standards of instruction or achievement on the part of a class, for the development of more harmonized progress, or for an increase in teaching efficiency on the part of the teacher. Classroom diagnosis is a process of class adjustment.

Educational diagnosis may also be concerned with the school. When the school is the focus of attention, it consists in a relationship between the administration of a school system and the various schools of which the system is composed. Here the technics of diagnosis are used to develop harmony within the system and to standardize, or fix, the levels of instruction, the rate of instruction, the classification of pupils, or the development of teaching. School diagnosis is a process of adjustment of the schools themselves to the needs of the system of which they are a part.

Range of individual diagnosis. Individual diagnosis, as has been stated, is concerned with every phase of the learning process and with every element which may affect a pupil in learning. It is, of course, impossible for anyone to analyze for any pupil all the possible factors which might affect his learning; but there are many factors which are possibly not obviously a part of the learning process which frequently act to retard the learning, and which must be located and remedied before learning can take place as it should. This implies that educational diagnosis of individuals should be largely confined to those pupils who are in greatest need of help, rather than to those whose work is satisfactory.

With this broad range of possibility, even though the actual effort in diagnosis is confined to a few individuals who need it most, the technics for the location of whom will be described later, the process of diagnosis should cover the range in at least as orderly and systematic a fashion as possible. The field of diagnosis can be conveniently divided

into three general parts: first, the physical conditions of the individual which impede learning; second, the mental conditions which impede learning; third, the instructional conditions which impede learning. The order of analysis as given is the order of primacy of investigation. If there are physical conditions which impede learning, these must be corrected or accounted for before it is possible to weight the influence of mentality; similarly, little of value can be determined in the correction of instructional difficulties until the mental condition is known and discounted in the final analysis.

Diagnosis of physical difficulties. When there are physical difficulties present in a pupil at the time of taking a test, it is possible for them to prevent him from accomplishing as much as he normally might. Thus, when a pupil is selected for educational diagnosis of his difficulties because of his relative low standing upon a test, the first element to be questioned is whether the test reveals true difficulties in learning or a reaction due to physical defects.

The extent to which physical defects or physical difficulties can enter to influence the accomplishment of pupils is not known except in certain specialized instances, such as cases of deafness and the like, but it is known that the health of a pupil does have an effect upon his work. When the health is poor the pupil's school work will probably be affected adversely. Within the past few years the health of pupils has become one of the special cares of the school-teacher, and the school-teacher has become one of the more important persons responsible for the maintenance of health. Aside from the programs of personal hygiene which the school-teacher has been expected to foster with his pupils, he is increasingly being held responsible for the general examination of his pupils for physical deficiencies and for the report upon them to the proper health officials for remedy or treatment.

Deafness, adenoids, diseased tonsils, faulty posture, defective eyesight, excessive nervousness, malnutrition, and the like are all responsible for many of the difficulties experienced by pupils in learning, and are as much within the province of the teacher to report for correction or treatment as are evidences of any of the infectious diseases which may break out among pupils.

Dr. Thomas D. Wood, Professor of Health Education in Teachers College, Columbia University, has charted for the use of school-teachers the general defects or disorders of school children which it is within their province to consider when making an analysis of the physical health of their pupils. These are defects or disorders which might influence the learning of the pupils adversely and which are therefore of importance for the teacher to know when beginning diagnostic work with his pupils. This chart groups the physical disorders or defects under three headings. In the first group are listed indications which teachers should be trained to notice and report upon because they point to disorders that may have serious consequences and require delicate adjustments. Such indications are nausea, fever, sore throat, swollen glands, and the like. In the second group are included signs of abnormality which point to more chronic disorders than are included in the first group and which should be remedied early. These include signs of disorders of the nose, throat, and ear, of the eye, of the teeth, of the feet or posture, and nervous disorders. In the third group are included indications of disturbance which are important in connection with other signs of physical disorder, such as general disorders including deficient or abnormal weight, lassitude, fatigue, low endurance, and the like, and defects of the feet and legs and defective movements. This chart, because it is so easy for the teacher to use and because it is so important in helping to determine physical causes of difficulty

which might influence school work, is quoted in full at the conclusion of this chapter.

Since, moreover, the ability to recognize normal conditions of health is important in the location of abnormal conditions, a further statement is quoted at the conclusion of this chapter, and in the bibliography there is listed an important scale of normal-health goals.

When a teacher is satisfied that a pupil's physical condition is normal and has not materially contributed to his lack of achievement or effort, and that there is no physical condition which if remedied would at the same time restore him to a normal level of achievement or effort, the second step in the process of diagnosis can be taken.

Diagnosis of mental difficulties. Variations in mentality are responsible for many of the seeming difficulties which pupils experience in learning. It is known that if a pupil has a so-called low mentality it is more difficult for him to learn than for pupils with so-called higher levels of mentality. Mentality, then, has become practically synonymous with ability, and has frequently been considered as the prime necessity for ability. It is possibly true that this is the case, but every teacher can point to pupils who because of greater effort have been able to achieve more than pupils of greater mentality. The relation of ability to achievement will be discussed later; it merely remains to be pointed out here that whereas mentality is undoubtedly a criterion of certain normal limits of achievement, in general pupils of lesser mentality achieve more proportionally than do those of greater. When this is the case the determination of mentality merely points to possibilities and gives no real explanation of realities.

The teacher should also bear in mind that the determination of mentality carries with it no particular explanation of success or failure in specific items. Mentality in general in-

fluences the rate of learning in far greater proportion than it does the kind of learning that can take place. This is more true of the range of mentality which is found in the public schools than is popularly imagined. It is not usually true that certain pupils cannot learn arithmetic at all, but rather that they learn more slowly than do normal pupils. The emphasis, then, which the teacher should place upon an interpretation of mentality should be in terms more largely of rate of learning than of kind of learning.

Mentality, as has been pointed out in Chapter X, may be satisfactorily determined by means of group intelligence, or ability, tests. When the intelligence, or ability, of a pupil whose educational processes are under scrutiny has been so determined, it is possible to begin the preliminary diagnosis of his difficulties in learning.

Preliminary diagnosis of difficulties in learning. The preliminary diagnosis of the difficulties of pupils in learning consists, first, in the location of the pupils in need of diagnostic help and, secondly, in the definition of the element or elements in which they need help most. The difficulties are to be found largely in terms of subject-matter achievement, in the rate of learning, or in the quality of effort which has been expended. In any case, the first elements consist in the location of the pupils who show deficiencies in achievement, those whose rate of learning is below normal, and those whose effort is below par, since these are the ones who need help most.

The necessary facts for determining these elements or individuals are revealed in an analysis of the results of tests. For this purpose tests can be given and scored as already described, or else, if there is data available such as that included in the classification-data sheets, that data can be utilized. It is for this reason that copies of test results should be given to a teacher.

1. *Determination of status in subject achievement.* If it is assumed that a test or tests have been given to a group of pupils and that the raw scores have been determined, the process of diagnosis can be described as in the following sections. If the results of tests in terms of T scores, G scores, age scores, and the like have been used in a classification system, it should be understood that the preliminary steps given below can be much shortened.

STEP 1. *Transmuting raw scores into achievement measures*

For the determination of subject achievement and the location of pupils in need of diagnostic help, the first step is to transmute the raw scores on the test into achievement measures. In the scales which have been described the raw scores should be transmuted into either age scores, grade scores, or T scores. If age-grade or T-grade tables have been constructed, it is always a good policy to list the scores for the pupils in terms of both grade units and T or age units.

For the T scale this process has been illustrated for the fifth grade in Chapter IX, and the portion of the process there described should be used for this part of the diagnosis, assuming that it has not been done previously for purposes of classification. The pupils of a fifth grade and their raw T and G scores in arithmetic are listed as in Table LXXXI. The data are taken from columns 2, 6, 10, and 15 of Fig. 94. If the classification data were not available, they could be found by giving a test and scoring it, arranging the pupils in alphabetical order, listing the raw scores, turning the raw scores into T scores according to the methods previously described, and finally turning the raw scores into G scores.

STEP 2. *Determining achievement norms*

The second step is the determination of the achievement norms. This particular test in arithmetic was given in the early part of May, and the pupils had completed eight months

TABLE LXXXI. ACHIEVEMENT SCORES IN ARITHMETIC FOR A FIFTH GRADE

No.	PUPIL	AR.S.	T_a	G_a
1	A	16	34	4.7
2	B	21	41	5.8
3	C	26	50	7.2
4	D	24	46	6.5
5	E	24	46	6.5
6	F	21	41	5.8
7	G	24	46	6.5
8	H	33	70	11.6
9	I	25	48	6.8
10	J	25	48	6.8
11	K	22	42	6.0
12	L	26	50	7.2
13	M	21	41	5.8
14	N	16	34	4.7
15	O	21	41	5.8
16	P	22	42	6.0
17	Q	25	48	6.8
18	R	30	59	9.3
19	S	26	46	6.5
20	T	23	44	6.2
21	U	26	50	7.2
22	V	21	41	5.8
23	W	23	44	6.2
24	X	28	54	8.0
25	Y	28	54	8.0
26	Z	21	41	5.8
27	AA	21	41	5.8
28	BB	21	41	5.8
29	CC	22	42	6.0
30	DD	29	56	8.6
31	EE	14	32	4.2
32	FF	27	51	7.5
33	GG	26	50	7.2
34	HH	27	51	7.5
35	II	21	41	5.8
36	JJ	19	38	5.3
37	KK	24	46	6.5

of the school year when the test was given; the grade norm is therefore the achievement of pupils at the end of the eighth month of fifth grade, which is 5.8. The T norm can be found by consulting the T–G table for the test and finding the T score for a grade norm of 5.8. In this arithmetic test the grade norm is 5.8, and the T norm as taken from the T–G table is 41.

If the age scale were being used, the process up to this point would mean turning the raw scores on the tests into arithmetic ages, and the ages into grade scores, as shown in Table LXXXII.

STEP 3. *Comparing achievement scores with grade norm*

The third step in the determination of subject achievement and the location of pupils in need of diagnostic help consists in a comparison of either the T or age score and the grade score with the norms which have been determined, so as to find the pupil or pupils who are lowest in achievement. In Table LXXXI Pupil 31 has the lowest T score, 32, and Pupils 1 and 14 have the next lowest, both 34. By comparison with the grade norm, 5.8, Pupil 31 is one grade and six months below the normal status for the grade, and Pupils 1 and 14 are one grade and one month below that status. This necessitates an analysis of the work of these pupils in their arithmetic to determine the difficulties which they have had and, if possible, the causes which resulted in the difficulties.

STEP 4. *Testing analysis for selected pupils*

The fourth step in the determination of subject achievement consists in an analysis of the results of the test which was given in arithmetic to discover the errors these pupils have made. This should reveal the weaknesses back of the failure of the pupils to achieve more than they did, and give the teacher an opportunity to bring to these pupils the special help of which they are evidently in need.

TABLE LXXXII. ACHIEVEMENT MEASURES IN ARITHMETIC FOR A FIFTH GRADE

No.	Pupil	Ar. S.	Ar. A.	G_a
1	A	16	130	4.7
2	B	21	144	5.8
3	C	26	161	7.2
4	D	24	153	6.5
5	E	24	153	6.5
6	F	21	144	5.8
7	G	24	153	6.5
8	H	33	196	11.6
9	I	25	156	6.8
10	J	25	156	6.8
11	K	22	146	6.0
12	L	26	161	7.2
13	M	21	144	5.8
14	N	16	130	4.7
15	O	21	144	5.8
16	P	22	146	6.0
17	Q	25	156	6.8
18	R	30	180	9.3
19	S	26	161	6.5
20	T	23	149	6.2
21	U	26	161	7.2
22	V	21	144	5.8
23	W	23	149	6.2
24	X	28	171	8.0
25	Y	28	171	8.0
26	Z	21	144	5.8
27	AA	21	144	5.8
28	BB	21	144	5.8
29	CC	22	146	6.0
30	DD	29	176	8.6
31	EE	14	124	4.2
32	FF	27	165	7.5
33	GG	26	161	7.2
34	HH	27	165	7.5
35	II	21	144	5.8
36	JJ	19	138	5.3
37	KK	24	153	6.5

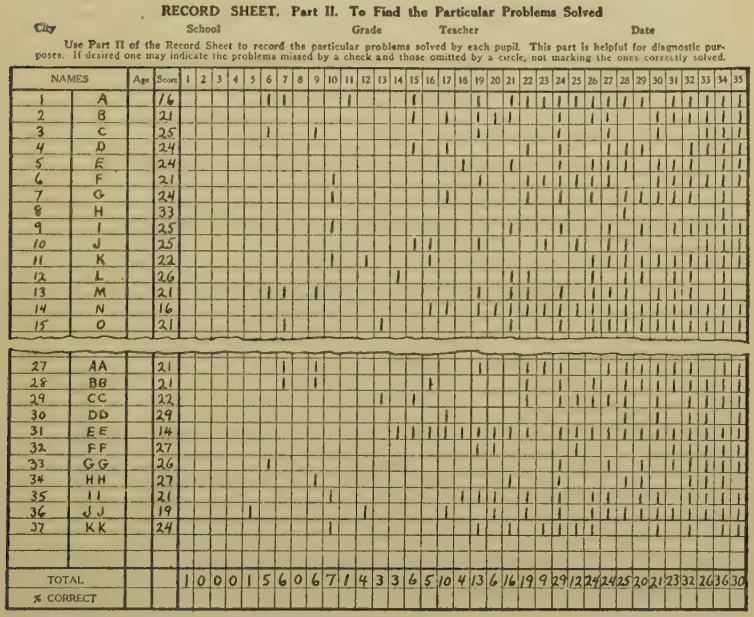
The process of location of the errors is one to determine what arithmetical errors these pupils made that were not characteristic of all the pupils. It is not expected that in a test which can be given to all the pupils in an elementary school above the third grade, as in the case of this test, the pupils in the fifth grade who are normal will not have errors, and an analysis of the test results will show there are portions of the test which few of the pupils could do successfully; but it is expected that the three pupils selected for special treatment have made errors such as should not have been made by pupils in the fifth grade. Such errors are not characteristic of the class as a whole, and they can be found in the following manner.

Among the materials furnished with the Woody-McCall Arithmetic Test is an analysis sheet. This sheet contains spaces for recording the errors made by each pupil, and also columns to show the total number of errors in the class in each of the problems of the test. It is a simple matter to fill in this sheet for the class from the original test papers and to add the number of errors for each column. When this has been done, the analysis for fifth grade has the appearance of Fig. 98. The figure shows that the pupils in this class made errors in the various problems of the test as shown at the foot of the columns.

From Fig. 98 it can be seen that no problem was too difficult for some pupil in the class, and that problems 2, 3, 4, and 8 were solved correctly by all the class. The remaining problems were solved with varying success.

A comparison of the problems failed of solution by Pupils 1, 14, and 31 with the solutions failed by the class as a whole shows that Pupil 1 failed to solve problems 6, 7, 11, 15, 19, 21, 23, and 25, which were failed in by less than half the class. Of these problems 6, 7, 11, and 15 were failed in by six pupils or less. These latter problems probably represent

characteristic errors of Pupil 1 and are errors which are not characteristic of the remainder of his class group. Reference to the original test sheet will show that these problems are easy subtraction, easy addition, easy division, and easy



RECORD SHEET. Part II. To Find the Particular Problems Solved

City School Grade Teacher Date

Use Part II of the Record Sheet to record the particular problems solved by each pupil. This part is helpful for diagnostic purposes. If desired one may indicate the problems missed by a check and those omitted by a circle, not marking the ones correctly solved.

	NAMES	Age	Score	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
1	A		16																																			
2	B		21																																			
3	C		25																																			
4	D		24																																			
5	E		24																																			
6	F		21																																			
7	G		24																																			
8	H		33																																			
9	I		25																																			
10	J		25																																			
11	K		22																																			
12	L		26																																			
13	M		21																																			
14	N		16																																			
15	O		21																																			
27	AA		21																																			
28	BB		21																																			
29	CC		22																																			
30	DD		29																																			
31	EE		14																																			
32	FF		27																																			
33	GG		26																																			
34	HH		27																																			
35	II		21																																			
36	JJ		19																																			
37	KK		24																																			
	TOTAL			1	0	0	0	1	5	6	0	6	7	1	4	3	3	6	5	10	4	13	6	16	19	9	29	12	24	24	25	20	21	23	32	26	36	30
	% CORRECT																																					

FIG. 98. Analysis of errors in arithmetic. (Portion of class record sheet for Woody-McCall Arithmetic Test)

multiplication. This indicates that Pupil 1 is probably weak in his ability to accomplish simple arithmetic computation.

Carrying out the same procedure with the work of Pupil 14 shows that, with the exception of problem 19, she has successfully completed all that she attempted. The analysis of the work of Pupil 31 shows that her characteristic errors are similar to those of Pupil 14, but that even less was accomplished.

Fig. 99 shows the errors of these three pupils, together with the total errors for the class, and makes somewhat easier the analysis which has been summarized above.

It is suggested by the foregoing analysis that the difficulty of Pupil 1 lies in lack of knowledge of fundamental arithmetical processes and that Pupils 14 and 31 are extremely slow in accomplishment.

Before a definite program of remedial work can be attempted, the question arises whether the work of these three pupils has been delayed through a slowness of work due to a

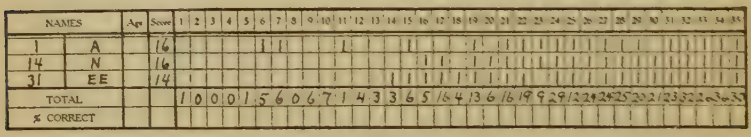


FIG. 99. Comparison of three pupils with class totals. (Taken from data of Fig. 98)

lesser ability than that of other pupils, or whether the deficiency is due to a lack of effort. If the deficiency is due to the fact that the normal rate of learning of these pupils or any of them is less than that of the other pupils, it is suggested that the pupils have probably been promoted, owing to an ideal for promotion of one grade a year, beyond the point which they should have reached. If these pupils, or any of them, are the equivalent of the other class members in their ability or rate of learning, then the deficiency may be due to a lack of effort, which should suggest a different treatment of the pupils so as to stimulate them to increase their effort.

2. *Determination of status in rate of learning.* As was shown in Chapter XI, the rate of learning of pupils varies greatly. Some pupils can learn more rapidly than others, and the rate of learning is shown by a relationship between their chronological ages and their abilities. In the age scale this

rate of learning is shown by the achievement quotient, and in the T scale by the B score in the subject.

By means of the method described on page 282 the first step in determining the B score in arithmetic for Pupils 1, 14, and 31 is to find for each of them the B correction for their chronological ages by means of Dr. McCall's B-Correction Table.[1] Table LXXXIII shows the numbers of the three pupils, their chronological ages, and their B corrections as taken from the table.

TABLE LXXXIII. B CORRECTIONS FOR THREE PUPILS

PUPIL	CA	BCR.
1	11–5	+4
14	13–6	−4
31	10–7	+8

When the B corrections have been determined they should be added (regarding signs) to the T scores in arithmetic (T_a) to give the B scores in arithmetic (B_a). Table LXXXIV shows the numbers of the three pupils under consideration, their T scores in arithmetic, the B corrections as transferred from Table LXXXIII, and the B score in arithmetic (B_a), which is the measure of rate desired.

TABLE LXXXIV. B_a SCORES OF PUPILS OF TABLE LXXXIII

PUPIL	T_a	BCR.	B_a
1	34	+4	38
14	34	−4	30
31	32	+8	40

If the age-grade scale were to be used for determining the measures in the status of these pupils in rate of learning, the procedure would be to determine the achievement quotients in arithmetic as described on page 280.

An interpretation of the rates of learning of these three pupils, as determined by the data in Table LXXXIV, by the T scale is as follows: Pupil 1, by Table LXXXIV, is 12 T units below the average in rate of learning for pupils of his age. It is evident that he is slow; but Pupil 14 is even slower,

[1] See page 275.

since she is 20 T units below the average in rate of learning for pupils of her age. Pupil 31, even though her T score is below that of the other two pupils, is better than either of the others in her rate of learning, being 10 T units below the average of pupils of her age.

All three pupils show a rate of learning which is below average, just as they showed a learning achievement which was below average. The question that arises is whether these pupils are learning as fast as they ought, considering their abilities, or whether they should be encouraged to learn more rapidly.

3. *Determination of status in effort in learning.* The determination of the status of these three pupils in their effort in learning will complete the preliminary diagnosis of their work. In the T scale this status is determined according to the procedure described on page 300. The technic requires that an intelligence test shall have been given to the pupils, for comparison with the subject-matter ratings in achievement.

The steps in the determination of the status in effort in learning, by the T scale, are first to find a T score in intelligence for each of the pupils by transmuting the raw scores on an intelligence test into T scores (T_i). These T scores must then be transmuted into B scores in intelligence (B_i) by adding (regarding signs) the B correction as found for the chronological ages of the pupils from Table XLVI. When this has been done, the next step is to apply the effort-score formula (see page 301).

$$(B_e + 50) - B_i = F_e.$$

For Pupil 1 the calculation would be as follows: his T score in arithmetic is 34, his B correction is + 4, his T score in intelligence is 39, and his B correction is + 4.

$$34\ (T_a) + 4\ (BCr) = 38\ (B_a),$$
$$39\ (T_i) + 4\ (BCr) = 43\ (B_i),$$
$$[38\ (B_a) + 50] - 43\ (B_i) = 45\ (F_a).$$

The necessary data for all the pupils is shown in Table LXXXV, where the calculations are given in transmutation from the T scores throughout. If desired, the T scores can be used if B scores are not wanted for other purposes. The data of Table LXXXV are taken from Fig. 94.

TABLE LXXXV. F SCORES FOR THREE PUPILS IN ARITHMETIC

PUPIL	CA	T_a	BCR	B_a	T_i	B_i	F_a
1	11–5	34	+4	38	39	43	45
14	13–6	34	−4	30	41	37	43
31	10–7	32	+8	40	46	54	36

When the age scale is used to determine the status in effort in learning, the procedure is to find the achievement ratios as described on page 296.

Summary of preliminary diagnosis. Assuming that the preliminary diagnosis has shown no serious physical defect in any of the three pupils under consideration, and that the measure of mentality has been made and used as shown in the analysis of their learning process, the preliminary diagnosis has revealed differing probabilities for the three pupils.

Pupil 1 has shown a difficulty in simple arithmetical procedures by being unable to do correctly such problems as the following:

Add
17
8

$3+1=$ $4+2=$

With such deficiencies as would result in these mistakes, it is evident that this pupil is deficient in fundamental number concepts.

So far as rate of learning is concerned, Pupil 1 is slower than normal pupils, and when his effort is measured he is even slower than his mentality would lead one to expect. It is evident that Pupil 1 should be expected to learn more slowly

than normal pupils, but it is also evident that he should be encouraged to learn faster than he does. To aid him to learn faster means, as will be shown, first to discover his real weaknesses and then to encourage him to overcome them.

Pupils 14 and 31 in achievement show a different characteristic from Pupil 1. In so far as these two pupils have attempted the test, they have worked it fairly correctly, but they have not done as much as is normal for pupils of their grade. Pupil 14, thirteen years and six months old, is overage for her grade, and her rate of learning is the slowest of the three pupils under consideration, but she also is not making an effort in her learning comparable with her ability. It is probable that she is working far too slowly, and a close examination of her methods of work will probably indicate faulty or time-consuming crutches. Finger-counting, short attention span, bad habits of work, and the like can be suspected; but slow comprehension is the most likely cause, and, as will be shown, this difficulty will be hard to overcome.

Pupil 31 is similar to Pupil 14 in general, save that her chronological age is more normal for the class she is in. The ability of this pupil is above normal for pupils of her age, yet both her rate of learning and her effort are below normal. There should be no reason for slowness in comprehension in this pupil, so that the most likely causes of her difficulties are to be found in faulty methods of work. Her effort is the lowest of the three, and it is therefore probable that she needs a special study of her difficulties in order to help her to achieve more nearly normal progress.

Intensive secondary diagnosis. The preliminary diagnosis of a school pupil does not reveal the causes of the difficulties of the pupil in learning. It does reveal in a better way than heretofore possible the probable location of those causes. Proceeding, therefore, from this preliminary diagnosis, to

gain a better understanding of the difficulties of the pupil by making closer inquiry into the cause of the difficulties, it is necessary to make a more intensive secondary diagnosis. This diagnosis is one which concerns the individual pupil so peculiarly that it is possible here only to outline the procedure to be followed.

The first purpose in making this diagnosis should be to locate, so far as may be, the causes of the pupil's difficulties in dealing with his subject matter, since this is the primary source through which his knowledge can be gained; a second purpose should be to locate the causes for his difficulties in rate of learning, if any, since, as has been pointed out, other factors than mere inherent ability can enter to influence this rate; a third purpose should be to locate the causes of his difficulties in effort, since these difficulties are sure to influence his learning.

Analysis for the location of difficulties in subject matter. Difficulties in subject matter are to a certain extent revealed through testing by achievement tests, as was shown in the discussion of preliminary diagnosis. Practically any test has some diagnostic value, but there are many tests now published which are designed almost exclusively for the diagnosis of difficulties in subject matter. Such tests as the Monroe Diagnostic Tests in Arithmetic and the Gates Tests in Reading have been mentioned. There are tests in many other fields of subject matter also especially designed for diagnostic use.

The purpose of these tests is to include within them a sufficient amount of material which also requires a wide variety of abilities or knowledge to deal with successfully, so that an analysis of the errors made by a pupil will show on the one hand the elements with which he can deal successfully, and on the other the elements in which he needs further knowledge or practice. In the direction manuals for these tests are

found specific directions for giving the tests and for analyzing the results obtained from them.

The teacher should remember that the fundamental purpose of the analysis of difficulties in subject matter is to guide him in future instruction, not merely to enable him to tell the pupil what is wrong.

Gates[1] reports a case study of a pupil deficient in ability to read, which shows a method of case analysis and of location of difficulties in learning to read. Included in this case study, which is reproduced below, is a summary sheet of the quantitative records obtained from tests used by Dr. Gates and available for the use of teachers.

Case 1. Deficiencies in Methods of Learning and Perceiving Words

The quantitative results for this case are given in the accompanying reproduction of the record sheet.

Case 1, a public school pupil, was near the end of Grade II when examined. He was then 9.7 years old. His mental age was 9.8 years which is about that of the average pupil of the same age. He had failed of promotion in the previous year because of his almost complete inability to read and was expected to fail again.

Reading achievements. Measured with the three Primary Tests, Case 1 showed ability to identify a half dozen words but was unable to read a sentence or paragraph. His difficulty was of long standing. He was believed to be incapable of learning to read.

Intelligence. The boy was given the entire series of tests in order to see whether his deficiencies were constitutional or due to acquired habits. In intelligence he was average; his Intelligence Quotient on the Stanford-Binet was 101. Such a score indicates, incidentally, that he can comprehend spoken, even if not printed, language as well as the average pupil. Two possible causes of the reading disability are thereby eliminated.

[1] Arthur I. Gates, *The Improvement of Reading*, pp. 317–324. The Macmillan Company, New York, 1927.

Diagnosis of Reading Quantitative Records

Name Case 1 Age 9.7 yrs. Birthday
School (Public) Room Grade 2.8 Date: April 10, 11, 1925

	Raw Score	Age Score	Grade Score
I. Intelligence			
1. Stanford-Binet . .		9.8	4.0
2.			
II. Primary Reading Tests			
1. Word Recognition	6	6.8	1.4
2. Phrase and Sentence Reading . .	2	6.7	1.2
3. Paragraphs; Directions	0		
III. Grades III–VIII Reading Tests			
A. General impression	0		
B. Predict outcome .	0		
C. Precise directions .	0		
D. Details	0		
IV. Oral Reading and Pronunciation			
1. Gray's Oral Passages	0		
2. Gates Word Pronunciation	6		
V. Phonetic Abilities			
A 1. Capital letters .	41		
A 2. Small letters . .	38		
A 3. Sound letters . .	18	8.3	2.8
A 4. Sound phonograms	32	7.9	2.4
A 5. Sound combined phonograms . .	4.5	7.2	1.7
B 1. Give letters for sounds	9.5	10.2	4.4
B 2. Give letters for syllables	7.0	8.9	3.2
B 3. Give letters for two-syllables . .	3.5	8.3	2.8
B 4. Spell words . .	5.0	7.2	1.7

	Raw Score	Age Score	Grade Score
VI. Visual Perception .			
A 1. Perception, figures	23	11.0	5.2
A 2. Perception, digits	35	9.3	3.6
A 3. Perception, words	22	7.3	1.8
B 1. Selection, figures	15	10.2	4.4
B 2. Selection, words	10	7.1	1.6
VII. Visual Analysis and Recognition			
1. Geometrical designs	7.0	9.8	4.0
2. Wordlike characters	6.0	8.6	3.0
VIII. Visual Memory Span			
1. Span for figures	4.0	9.5	3.8
IX. Auditory Functions			
1. Auditory acuity	O.K.		
2. Auditory discrimination	6.0	10.0	4.2
X. Auditory Memory Span			
1. Digits	6.0	10.4	4.5
2. Letters	5.0	9.0	3.3
3. Nonsense words .	4.0	9.2	3.5
4. Words	5.0	10.4	4.5
XI. Associative Learning			
A 1. Visual-visual .	7.0	9.1	3.4
A 2. Visual-visual .	8.0	10.6	4.7
A 3. Visual-visual .	5.0	7.7	2.2
B 1. Visual-auditory	7.0	9.3	3.6
B 2. Visual-auditory	8.0	7.8	2.4
B 3. Visual-auditory	5.0	7.3	1.8

Vision and visual perception. The scores for visual perception are shown under heading VI in the Record. Note that in speed and accuracy of perceiving geometrical figures (Tests A 1 and B 1) he secured a higher score than the average child of his age and intelligence. In the perception of digits (Test A 2), his attainment is lower but only a trifle low for his age. In perceiving words (Tests A 3 and B 2) he is far below the average ability.

These results indicate that he is not subject to any general deficiency in ability to perceive small visual items to a degree that could account for his failure to learn to read. They indicate, also, that he probably has no serious defects of the visual apparatus. He is, however, seriously deficient in ability to perceive rapidly and accurately printed words. When the deficiency in perception is confined to words, it is more likely to be due to unfavorable habits than to organic defects.

Visual analysis and recognition. The results of the two tests for visual analysis and recognition (Tests VII, 1 and 2) give no occasion for concern. Case 1 does as well as might be expected in the first, and somewhat less well in the second which is composed of characters similar in certain respects to printed words. The latter deficiency is not serious enough to account for the inability to learn to recognize words.

Memory span. Averaging the results for the visual memory span test (VIII, 1) and the auditory series for digits, letters, nonsense, and senseful words (X, 1, 2, 3, 4) we secure a composite age of 9.7 years, which, by a mere coincidence, is exactly equal to the pupil's mental age. The pupil's difficulty in reading, therefore, cannot be attributed to a poor immediate memory.

Auditory acuity and discrimination. Test IX, 1, showed Case 1 to have normal hearing and normal ability to discriminate and retain, momentarily, complex word-sounds. This fact eliminates auditory defects from the possible causes of disability in reading.

Associative learning. The first two tests in visual-visual association indicate approximately average ability; the third, which utilizes wordlike characters, shows a clear deficiency. Tests B2 and B3 of the visual-auditory series show a similar inferiority. In observing the work in these tests, the examiner was impressed with the pupil's clumsy technique especially when rather complex visual configurations were used. In the Recognition Test (Test VII, 2) the pupil labored distressingly in studying similar figures but earned a score only a year lower than his chronological age. In the learning tests, similar effort was expended, but, when the added requirement of associating the symbol with a picture or spoken word was undertaken, his achievements were two or more years below the norms for his age. To analyze complex figures,

then, appears to be somewhat of a difficulty; a difficulty sufficient to interfere considerably with associative learning of such symbols. The low score in the learning test was probably largely due to inefficiency in visual analysis of the material to be learned since in the tests composed of simpler figures, his learning scores were higher. It does not appear that this pupil is subject to defective visual or auditory memory, to deficient visual-visual or visual-auditory association brain tracts, to congenital alexia, and the like. He suffers more, probably, from deficient techniques of perceiving and learning wordlike characters.

Phonetic abilities. A pupil who cannot read and who can recognize but few words can scarcely be expected to excel in tests of phonetic ability. Case 1, nevertheless, had developed some skill in giving the sounds which letters represent and even of working out some of the phonograms by a laborious method. He could also, when given certain letter or syllable sounds, tell the letters which stood for them. He could spell orally about as many words as he could read. Although he could say his *a b c*'s, he made errors in reading some of the letters (Tests V, A1 and A2) and read some of the others so slowly that his time score fell below our norms. In these tests, the examiner felt that the pupil demonstrated about as much phonetic skill as might reasonably be expected of a pupil of so limited reading ability.

Methods of studying and perceiving words. The Gray's Oral Passages and the Gates Pronunciation Tests were applied in the manner outlined in Chapter VI. The pupil's difficulties in word recognition were so great that he received a zero score in Gray's Oral Test. In the Gates Pronunciation Test only six words were correctly pronounced. In his work in these two tests and in the tests of phonetic abilities it became apparent that the pupil's methods of observing and studying word-forms were extraordinarily inadequate. He had no consistent method. He observed one detail after another; sometimes progressing from left to right, sometimes in the reverse direction, sometimes beginning at the middle of the word and moving in either direction. He sought primarily minute characteristics and rarely observed the larger combinations of letters which comprise syllables and phonograms. Often he merely looked at one part of the word. Some-

times he named certain letters, sometimes he sounded them. The general configuration of the word was rarely observed, apparently, and there was scarcely a word that he could recognize after an instant's exposure.

No other deficiency was found except the clumsy technique in associative learning mentioned above. Although he was greatly discouraged, the pupil's desire to learn seemed genuine. He was not an unstable boy but he displayed considerable strain and nervousness when trying to learn or read words. That he appeared to be emotionally sensitive and somewhat "scatterbrained" in this work may be attributed to several years of chagrin and discomfort connected with it. In other tasks these symptoms were not apparent.

Diagnosis. The final diagnosis was that while Case 1 showed no deficiencies in the capacities required for learning to read he had acquired techniques of word study and perception which practically prohibited progress. The word features which he laboriously studied were mainly those which are not useful in distinguishing one word from many others. His methods of observation, moreover, were so detailed and irregular that only rarely could he repeat the analysis sufficiently to recognize the same word-form. His was a real case of failing to see the forest because of the trees.

Remedial instruction. Case 1 was first given, by a competent teacher working under the writer's advice, twelve periods, from 30 to 45 minutes each, of drill with materials of the type outlined in Chapter VII (pages 161–166). The teacher also gave instruction of the types outlined in the last section of Chapter VII designed to call attention to the more significant features of words. Although progress was made, the pupil had great difficulty in learning to perceive and analyze the word-forms in the new ways. Accordingly, the writing method as outlined in Chapter XII was applied vigorously for ten days in combination with the other practices. In the middle of this work, the pupil seemed rather suddenly to grasp the method. He began to learn words and to remember them. Seeing his progress, the boy quickly developed an eagerness for learning and reading words that was almost pathetic. He studied words during every spare minute; he tried his skill on papers, posters, signs — whenever he saw a word he tried to

figure it out and to remember it. His progress became amazingly rapid. He learned nearly a hundred words during the first week after the training took effect. After three more weeks he could read First Readers with ease; a month later, his reading on standard tests equaled that of the average pupil at Grade 2.5. After approximately three months of miscellaneous reading during the summer vacation, Case 1 earned on the Gates Silent Reading Tests the following scores:

A. General Impression; Reading Age, 10.0; Reading Grade, 4.2.
B. Predict Outcome; Reading Age, 9.7; Reading Grade, 3.9.
C. Precise Directions; Reading Age, 9.8; Reading Grade, 4.0.
D. Details; Reading Age, 10.1; Reading Grade, 4.3.

At this time the pupil's age was 10.3 years and his mental age about 10.4 years. During the period of six months he had developed from practically a nonreader to nearly an average reader for his age and intelligence.

Analysis for the location of difficulties in rate of learning. Difficulties in potential rate of learning, it will be remembered, are indicated by low intelligence quotients or low B scores in intelligence. Difficulties in previous rate of learning are indicated by low achievement quotients or by low B scores in education.

As has been already pointed out, the intelligence quotient and also the B score in intelligence are relatively permanent scores. They are based on native ability and are therefore, when carefully determined, fairly indicative of a relatively unchanging characteristic. When these qualities related to native ability are low, it becomes difficult to deal with the individuals concerned on the usual school levels. Their rate of learning and of comprehension is slower than normal and must be so recognized. It is a fact which must be accepted and discounted in teaching. Teachers of such pupils should place their main effort upon keeping the effort scores normal and upon exercising unfailing patience and understanding.

When previous rate of learning is under consideration, since many factors other than inherent ability can enter and influence the learning, it is important that a careful analysis of the probable causes of difficulty be made. When the achievement quotient is approximately 100 or when the B score in education is approximately 50, the indications are that the pupil is progressing normally for his age. Thus, regardless of what grade a pupil may be in, these two scores reveal what the school has done for the pupil, and a consideration of these scores in conjunction with the intelligence quotient or the B score in intelligence is of service.

When the previous rate of learning is slower than the potential rate of learning, there are several possible causes which should be investigated. Among them are these:

1. *Improper classification.* If a pupil has been consistently classified with pupils considerably below him in ability and in a position where he could without great effort do satisfactory work relatively to these pupils, the achievement quotient or the B_e is likely to be low. The remedy in such a case is to advance a pupil as rapidly as possible so that he can eventually be placed in a position where his abilities will be challenged.

This situation is more common among school pupils than is usually suspected, and, as indicated, is more prevalent among the brighter pupils, since there is about ten times as much retardation in the average school as there is acceleration. A complete reclassification of the pupils in a school will do much to correct this difficulty.

2. *Difficulties in subject matter.* Some outstanding difficulty in subject matter may act to prevent a pupil from having as high an achievement quotient or B score in education as he ought to have. In such cases an analysis such as was suggested on page 386 can help to locate and remove the difficulty. The effects of the diagnosis of the pupil cited by

Gates is a splendid illustration of the result which can be obtained through removing a specialized form of difficulty of this nature.

3. *Lack of effort.* This phase of the matter will be discussed later in this chapter. As was pointed out in considering improper classification as a cause, lack of effort may be a fundamental cause of a lower rate of learning than normal.

4. *Lack of uniform growth.* When a pupil actively dislikes some particular form of learning, as is occasionally the case, and thereby neglects it as much as possible because of its disagreeable nature to him, a low achievement quotient or B score is a natural and understandable result. Such a dislike may be deep-seated or only superficial. When it is deep-seated, the cause is likely to be found in some difficulty in subject matter which if corrected will bring a different and better feeling toward the work. Most things that a pupil can do well and take pride in doing well are pleasurable activities. It is when a thing is done poorly that it is burdensome. When the aversion is superficial the difficulty is probably emotional and is likely to disappear with time or with a change of emphasis on the part of the teacher. It is not unusual to find pupils, for example, who read avidly in some types of subject matter, but who are intrinsically uninterested in other types and therefore read as little of them as possible. What is true in reading is equally true in arithmetic and the like. In cases of this type the work of the teacher is so to choose the materials or the method of teaching as to sustain the interest and accelerate the learning.

When the previous rate of learning is more rapid than the potential rate of learning, there are several causes which should be investigated. This is a less usual condition than the one discussed above, but it is one which results in greater danger and misery for the pupil and is therefore one to

investigate. Some of the causes for this state of learning may be among the following:

1. *Improper classification.* In this case the pupil is likely to be classified among pupils considerably his superiors both in ability and accomplishment. Assuming that these pupils are more likely than he to be rightly classified, the result is that the pupil lacks the background for the accomplishment which is wrongly expected of him, and is incapable of coping with his fellows. The tendency therefore, as shown in the next two discussions, is for undue pressure to be placed upon him or undue effort to be expected from him.

2. *Undue pressures.* When a pupil is rightly classified, but especially when he is wrongly classified, it is possible for undue pressures toward study or learning to be made on him. These pressures may be made either at home or at school, and frequently in both. It is characteristic of teachers to attempt to raise the standards of their classroom achievement by placing pressures, stimulations, encouragements, and the like on the less gifted pupils. It is also characteristic, and in the same vein, for them to judge the less capable pupils with greater severity than those more capable. It is this condition which results in the greater proportion of retardation in schools than of acceleration. The standard of classroom work which is set by the average of the class or by the average expectations of the curriculum is easily met by the more capable, but met with far greater difficulty by the less capable. Since this standard becomes the first consideration of the teacher, he feels it his duty to bring the entire class to it and therefore exerts proportionally far more effort with those below the average of the class than with those above it. Inevitably the greater effort (since the varying potential rates of learning of the individual pupils have been lost to sight) results in undue and unwarranted progress on the part of the less capable pupils. A proper consideration of the varying

potential rates of learning of pupils will result in the necessary corrective for the difficulty.

3. *Exceptional effort.* This phase will be more fully discussed later. Exceptional effort, however brought about, or whatever the cause, is in itself a frequent basis for a higher previous rate of learning than is wise to encourage.

4. *Lack of uniform development.* It is not true, necessarily, that an overdevelopment in one phase of school instruction results in an underdevelopment in some other phase which is also measurable.[1] It would seem clear, however, that an undue development of the pupil's interests in one field is likely to react to the detriment of some other phase. Such a condition occurs on occasion when a pupil becomes excessively interested in some field outside the school curriculum which tends to absorb his time outside of school and his thoughts when he is in school. This has been true in many instances in terms of radio or airplanes. The correction for such a condition is the inclusion of the distracting interest within the life of the school and its correlation with the desired school activities, so that each can be enabled to receive its proper or desirable emphasis.

Analysis for the location of difficulties in effort. As was shown on page 294, the degree of effort made by a pupil consists in a relation between his native ability as expressed in his mental age or his T or B scores in intelligence and his actual achievement as expressed either in subject ages or in T or B scores in education.

In the age scale an achievement ratio of approximately 100 indicates a degree of effort normal for the individual, and the same fact is revealed in the T scale through F scores of approximately 50. When the achievement ratios are either greater or less than 100, or when the F scores are either

[1] Compare B. R. Buckingham's *Research for Teachers* (Silver, Burdett and Company, New York, 1926), p. 201.

greater or less than 50, an investigation should be made to determine the causes for the variation from normal. When effort scores are low, there may be a variety of possible causes, among which are the following:

1. *Lack of proper recognition of the ability of the pupil.* When the true ability of the pupil is unrecognized, as is the case with many superior pupils, the effort scores are likely to be low. The judgment of a teacher with respect to the ability of a pupil is frequently in error. Brightness of eye, tensity of attention, apparent rapidity of response, and the like are used as criteria of superior ability, and, conversely, general unattractiveness in expression, laxity in attentiveness, and lack of eagerness in response are only too frequently considered as criteria of dullness. With superior pupils, however, with whom the work in hand is simple, and to whom the elementary character of the work is inevitably tedious and unnecessary, active attention is unnecessary and quickness of response is not stimulated. Such pupils, if they be normal, are mischievous, are inattentive, are playing or drawing or thinking of something which does challenge their abilities, and are therefore classed as lazy or indifferent. The mere recognition of the pupil's true ability, and insistence upon worthy activity is the surest corrective.

2. *Lack of challenge in the work.* Just as lack of recognition of superior ability is likely to result in indifference, so lack of challenge in the work is likely to result in a feeling of inferiority. When the effort score is low and the ability score higher than that for other pupils in his class, the probability is that the challenge of the work to the pupil is practically nil. Attention may here be called to the case of Pupil CC shown in Fig. 94, with an F score of 33. This child, only ten years and six months old, although in the fifth grade in May and with May sixth-grade achievement, has an ability equivalent to that of ninth-grade pupils (the third year of

junior high school). Had his ability been recognized earlier in his schooling, and had he been placed more nearly where the work would have challenged him, his effort score would unquestionably have been much higher. His teacher, and justly so, adversely criticizes his effort, attentiveness, playing in school, and the like but places the onus *upon the pupil* and not upon his improper classification, where it properly belongs. When the pupil was questioned with respect to his effort, and particularly with respect to the annoying activities, he said: "I pay attention when I need to. When there are new things I pay attention all right; but most times it's the same old stuff, and I just close up my ears and think of something interesting [sic] until the other kids have caught up!"

When the effort score and the ability score are both low, it is probably indicative of the fact that the ability of the pupil has been underestimated and that therefore even less was expected of him than he actually could accomplish. The pupil needs more stimulation than he has had.

3. *Lack of adaptation of the work to the pupils.* A teacher who consistently underestimates the ability of his pupils is likely to find consistently low effort scores for the large majority of them. This may result either from an underestimation by the teacher or from a consistently low classification of the pupils. In either event a better understanding of the ability of the pupils, a change in the type or method of work, or a better adaptation of the work to the pupils will be the proper corrective. In the seventh-grade group listed in Chapter XV,[1] if the F scores (or achievement ratios) in reading are calculated, it will be found that eleven of the fifteen pupils have F scores below 50, a probable cause being an underestimation, on the part of the teacher, of the pupils' potential ability to read.

[1] See page 415, Table LXXXVIII, and page 421, Table XCI.

4. *Improper classification.* As implied above, a possible cause for low effort is likely to be found in the improper classification of the pupil, due to a lack of understanding of his real ability. In many cases a more rapid acceleration, in spite of the teacher's opinion if necessary, will be found to be the proper corrective. When effort scores are low, however, and ability scores are also lower than normal for the group in which a pupil is located, acceleration will be a mistake rather than a help.

When effort scores are high, there may be a number of possible causes, among which the following may be suggestive to the teacher:

1. *Overstimulation.* Urged by the teacher in school or by the parents at home, or by both, beyond their normal rates of progress, some pupils are enabled to exceed normal effort scores. This has previously been commented upon. Just what effect such overstimulation may have upon a pupil is doubtful. It is probable, however, that the reaction is more likely to be one connected with the nervous system than with the educational development. A pupil should not be expected to do better in any field than his ability would warrant, and if confidence is placed in the measurement of his ability it is probably unwise to stimulate him beyond that point. Parents, however, are ambitious for their children, and the pressures placed upon advances in schooling by the school itself are taken most seriously by many pupils. The social stigma is marked, especially for pupils who are graded solely on the basis of absolute achievement (or achievement in terms of group placement), and "failure" is far too greatly emphasized.

The corrective for overstimulation is a better understanding of the pupil both on the parent's part and on the teacher's, coupled with a more reasonable attitude toward failure (see Chapter XVII). It is only fair to measure a part of the pupil's

success in terms of his own ability and to aid him to find his own place rather than to usurp the place, progress, and recognition of another. Every pupil should be able to enjoy the true recognition of his intrinsic worth in terms of his own accomplishments.

2. *Lack of proper recognition of limitations in ability.* Closely associated with the overstimulation which was just discussed as a cause for high effort scores is that implied above in the lack of proper recognition of his limitations in ability. The discussion will be expanded in Chapter XVII,[1] but it should be stated here that a child should not be expected to do better than his best, and a true recognition of his limitations in ability will go far toward helping both teachers and parents to a better recognition of his accomplishment.

When there is doubt as to the validity of the measure of ability, a pupil should be retested. In the case of Pupil 9 of Tables LXXXVII and XC, there should be considerable doubt cast upon the results of the intelligence test. It seems rather improbable, although possible, that a pupil able to achieve a grade score of 13.8 in reading, of 8.6 in arithmetic, and of 6.2 in spelling (giving a combined effort score of 69) should have a grade score in intelligence of 5.2. In this case, before accepting the low ability rating as conclusive, the pupil should be retested.

3. *Unequal educational development.* A third possible cause of high effort scores may be found in an educational development which varies greatly in different fields. A pupil with a high effort rating in arithmetic and low effort ratings in other fields may be suspected of overworking in arithmetic at the expense of his other school work. Such a condition may be corrected by a more equable distribution of time, of incentives, or of interests. This is preëminently the work of the

[1] See page 456.

teacher, one which will do much to harmonize not alone the work of the pupil but also his position within the classroom.

The profile chart and the personal record of the pupil. Since ultimately the work of the teacher (and through the teacher, the work of tests and testing) is to bring about a more intelligent knowledge of the pupil, greater and greater emphasis

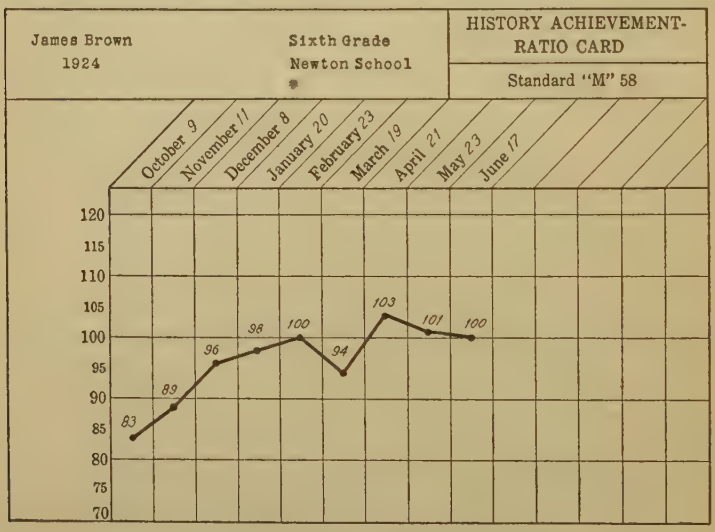


FIG. 100. A Progress Profile Chart

is being placed upon the record of the individual pupil. Many attempts have been made to devise schemes to make permanent records of the status and progress of pupils, not only in terms of the tests they have taken but also in terms of their social and other relationships. It is abundantly clear that the social background is a powerful factor in the development of a child, and that there are many interrelating factors of social and educational development. Frequently variations in one can be satisfactorily traced through a knowledge of the other.

One of the more successful means for bringing the entire record of a pupil to the attention of the teacher is the profile chart. This chart can take many different forms, depending upon the content of the chart and what is to be emphasized. It is possible, for example, to have such a chart indicate either status or progress. A progress profile chart for use in classroom tests is shown in Fig. 100.

A chart which shows the status of a pupil at a given time, and which also makes possible the listing of much information that will aid in the understanding of a pupil, made by Professor Paul R. Mort of Teachers College, Columbia University, is shown in Fig. 101. This is in the form of a card which can be filed and kept for reference. Some samples of this card as used in a school survey are shown on pages 402–404.

Chapter summary. In general, individual educational diagnosis should be confined largely to those pupils who are most in need of help. These pupils will be found either among the more capable in the class or among the less capable.

Preliminary to educational diagnosis, particularly in the case of backward or deficient pupils, should be a diagnosis of possible physical difficulties which might account for the backward condition. The probable physical difficulties which might be of adverse influence upon learning are summarized at the conclusion of this chapter, as arranged by Dr. Thomas D. Wood.

A second preliminary to educational diagnosis of pupils consists in a diagnosis of mental difficulties. Variations in mentality are responsible for much misunderstanding of the work of pupils and should be investigated before conclusive results can be obtained.

The preliminary educational diagnosis consists in the location of the pupils most in need of such help and in the location of such difficulties as the pupil or pupils may have. These difficulties can be revealed to a certain degree through

Pupil_____ Date_____ Grade_____ Boy_____ Girl_____

Address_____ Tel. No._____ Home Room_____ Parent_____

School_____ Teacher_____

Date of Birth_____ Height_____ Weight_____

Entrance grade in this school system_____ Entrance grade in this school_____

Grades skipped_____ Grades repeated_____ Half years spent in kindergarten_____ Does pupil expect to go to high school?_____ College?_____

What does he want to do when he finishes school?_____

Does he take music lessons outside of school?_____ Other lessons, if any, outside of school?_____

_____Member of the following school teams or other organizations:_____

_____ Outside organizations:_____

_____ What musical instruments?_____

What magazine does he read most?_____ How is out-of-school time spent?_____

_____ Is there a room in the home where he can study by himself?_____ What kind of music does he like best?_____ Has he a card for the public library?_____ Times each week spent at the movies?_____ What language is spoken in the home?_____ How many books in the home?_____

Specific Weaknesses_____

Physical Defects_____

Too mature socially for this group?_____ Too immature?_____ Does the program of this pupil vary from that of the class as a whole in any manner?_____

Has this pupil ever been in a class of other than average-ability pupils? (Give grade and nature of class and success)

_____

Variations from section program at any time in past (give grade, type of section, and nature of variation)_____

_____

Any other irregular features in this pupil's school history_____

Enrollment in this class_____ Type of class_____

FIG. 101 A. Mort's Pupil-Classification Card (obverse)

Draw a line through each item in which this pupil is one of the best five in his class; encircle each item in which this pupil is one of the poorest five in his class. Indicate scholarship in blanks:

1. General Ability _____ 5. Leadership _____ 9. Freehand Drawing _____ 13. Spelling _____
2. Mechanical Ability _____ 6. Reading _____ 10. Arithmetic _____ 14. Music _____
3. Attitude toward Work _____ 7. Nature Study and Science 11. History _____ 15. Shop Work _____
4. Health _____ 8. Language Usage _____ 12. Literature _____

Variation from Normal Weight _____
I.Q. _____ E.Q. _____ Average _____

Scores

Age Equivalents

GRADE

3 4 5 6 7 8 9 10

8 9 10 11 12 13 14 15 16

AGE

Chronological Age . . . . . _____ _____
Height . . . . . _____ _____
Mental Test . . . . . _____ _____
Achievement, Average . . . . . _____ _____
Reading, Paragraph Meaning . . . . . _____ _____
Reading, Sentence Meaning . . . . . _____ _____
Reading, Word Meaning . . . . . _____ _____
Arithmetic Computation . . . . . _____ _____
Arithmetic, Reasoning . . . . . _____ _____
Nature Study and Science . . . . . _____ _____
History and Literature . . . . . _____ _____
Language Usage . . . . . _____ _____
Spelling . . . . . _____ _____
Mechanical Aptitude . . . . . _____ _____

_____ Pupil's Record _____ Record of Section to which assigned.

Use vertical line to represent grade standard.

FIG. 101 B. Mort's Pupil-Classification Card (obverse) (continued)

Assigned Classification _____ _____ Term 19_____19_____
Grade_____ Home Room_____ Section_____ Type of Section_____
Variations from this Section_____

Preliminary Classifications

(1) Teacher's Recommendations: Grade_____ Type of Section_____
Special Needs_____

(2) Schedule Committee: Grade_____ Section_____ Type of Section_____
_____ Variations from this Section_____

(3) Reviewing Committee: Shortcomings in his classification_____

PUPIL-CLASSIFICATION CARD

Additional Information bearing upon the Classification of this Pupil

FIG. 101 C. Mort's Pupil-Classification Card (reverse)

an analysis of the results of standard tests, when the results are tabulated and analyzed to locate the characteristic errors. To do this, technics are described for the determination of the status of pupils in subject achievement as well as technics for determining the character of the errors which were made.

With the completion of the preliminary forms of educational diagnosis, intensive secondary diagnosis of the pupils who have been selected for further study can follow. This secondary diagnosis consists in the giving and analysis of diagnostic tests for the location and remedy of difficulties in subject matter, in the analysis of the pupil's work for the location of his difficulties in rate of learning and for the explanation and correction of the difficulties involved where abnormal effort ratings are revealed.

All these procedures eventually result in a more complete knowledge of a pupil than would otherwise be possible, and make judgments concerning the worth and character of the pupil's work more just, more enlightened, and more effectual.

Signs of Health Disorders and Physical Defects in School Children

[Arranged for teachers by Thomas D. Wood, M.D., Teachers College, Columbia University, New York City.[1]]

The following signs of disorder have been arranged in three groups for the use of teachers in detecting possible health and physical defects in children under their care.

Group I contains signs of disorder which teachers should be trained to notice and report to constituted authorities. Group II names signs of abnormality pointing to more chronic disorders which should be remedied early. Group III contains indications of disturbance which are important in connection with other signs of physical disorder.

[1] This table is a revision of the original table as given by Dr. Thomas D. Wood in *Transactions of the Fourth International Congress on School Hygiene*, Vol. IV, pp. 135–692.

GROUP I. *Indications of Health Disorders in Children which Teachers should be trained to notice and to report to Constituted Authorities*

Signs

- Nausea or vomiting
- Chill, convulsions (fits)
- Dizziness, faintness, or unusual pallor (alarming paleness of the face)
- Eruption (rash) of any kind
- Fever
- Running nose
- Red or running eyes
- Sore or inflamed throat
- Acutely swollen glands
- New cough
- Any distinct or disturbing change from usual appearance or conduct of child

The foregoing signs should be used by teachers as a basis for excluding pupils from school for the day or until the signs have disappeared or until the proper health officer has authorized the return of the pupil to school.

GROUP II. *Signs of Abnormality pointing to more Chronic Disorders which should be remedied early*

Signs

Disorders of nose, throat, ear, and organs of respiration:
- Mouth-breathing
- Loud breathing
- Nasal voice
- Catarrh
- Frequent colds
- Offensive breath
- Chronic cough
- Deafness
- Twitching of lips
- Headache

Eye disorders and defects:
- Headache
- Crossed eye
- Squinting
- Holding book too near face

Decayed teeth	Teeth defects
Crooked teeth	
Discoloration of teeth	
Offensive breath	

Inability to hold objects well	Nervous disorders
Spasmodic movements	
Twitching of eye, face, or any part of body	
Nail-biting	
Perverted tastes	
Sex disturbances	

Pain in feet	Defects of feet
Toeing markedly out	
Flatfoot gait	
Swelling, puffiness of feet	
Excessive perspiration of feet	

Unequal height of shoulders	Incorrect posture
Flat chest	
Round neck and shoulders	
Stooping	

GROUP III. *Indications of Disturbance which are Important in Connection with Other Signs of Physical Disorder*

Signs	
Deficient weight	Nutritional and general disorders
Pallor	
Lassitude	
Perverted tastes (food)	
Slow mentality	
Peculiar or faulty postures	
Underdevelopment	
Excessive fat	
Low endurance	
Disinclination to play	
Fatigue	

Pigeon-toed gait	Defects of feet and legs, and defective movements
Shuffling, inelastic walk	
Exaggerated knee action in walking	
Shifting from foot to foot	
Standing on outer edge of feet	
Standing on inner side of feet, heels turned out	
Locking knee	
Leaning against wall or desk	
Shoes run over at either side	
Wearing out soles asymmetrically	
Twitching of foot muscles	

SUPERVISION OF THE HEALTH OF THE PRE-SCHOOL CHILD: ITS RELATION TO THE EARNING CAPACITY OF THE WAGE-EARNER

[Read before the Section on Public Health and Industrial Hygiene, College of Physicians, Philadelphia, February 14, 1927.[1]]

. . . One reason for our previous failure to realize the poor physical condition of the average child was the fact that there have been few if any authoritative statements of exactly what is satisfactory health and development in childhood. Thus there has been no adequate norm against which to measure a child, and we have been satisfied with low standards, whereas it is quite possible to achieve much higher standards. More satisfactory signs of health are being determined all the time. Even a very simple statement will show to a certain extent the degree of physical health and development which may be acquired by a large majority of children.

Bright, clear eyes with normal focus and movements.
Unobstructed nasal breathing.
Clean, moist tongue.
Well-formed and well-enameled teeth.
Mucous membranes definitely pink.
Subcutaneous tissue plentiful.
Muscles firm and strong.

[1] Extract of article by Hugh Chaplin, M.D., appearing in the *New York State Journal of Medicine*, January 1, 1928.

Shoulders level; chest broad and deep, with good expansion.
Straight limbs and strong joints.
Arches of feet strong and limber.
Inner border of foot straight from heel to tip of great toe.
Weight suitable for height and age.
Good posture.
Alert, happy expression.
Prompt, efficient muscular coördination.
Bodily repose.
Physical and nervous endurance.

SELECTED BIBLIOGRAPHY

BROOKS, S. S. Improving Schools by Standardized Tests. Houghton Mifflin Company, Boston, 1922.

BUCKINGHAM, B. R. Research for Teachers, pp. 330–367. Silver, Burdett and Company, 1926.

GREENE, H. A., and JORGENSEN, A. N. The Use and Interpretation of Educational Tests, chaps. viii–xi. Longmans, Green & Co., New York, 1929.

KALLOM, A. W. "The Importance of Diagnosis in Educational Measurement," *Journal of Educational Psychology*, January, 1919, pp. 1–12.

MCCALL, W. A. How to Measure in Education, pp. 67–111. The Macmillan Company, New York, 1922.

MONROE, W. S., DEVOSS, J. C., and KELLY, F. J. Educational Tests and Measurements (Revised and Enlarged Edition), pp. 76–83. Houghton Mifflin Company, Boston, 1924.

MORT, P. R. The Individual Pupil. American Book Company, New York, 1928.

ODELL, C. W. Traditional Examinations and New-Type Tests, p. 36. The Century Co., New York, 1928.

PAULU, E. M. Diagnostic Testing and Remedial Teaching. D. C. Heath & Co., Boston, 1924.

WOOD, T. D., and LERRIGO, M. O. Health Behavior. Public School Publishing Company, Bloomington, Illinois, 1927.

CHAPTER XV

THE DIAGNOSIS OF CLASSES

Significance of class achievement. The work which is done by a class of pupils is not necessarily a complete reflection of the skill of the teacher, nor is it, so far as surface inspection is concerned, a measure of the pupils. A class whose total achievement may be below normal may yet have been carried further by a skillful teacher than one would have reason to expect, and another class with a total achievement above normal may yet be actually below the point where a more skillful teacher might have led them. A diagnosis of class achievement will aid in determining the attitude and feeling which the teacher, supervisor, or principal should have toward them.

Since there is wide variability in the abilities and potential and actual progress of pupils, no set curriculum can be rigidly maintained, either in terms of maximal attainment or in terms of rate. For this reason a diagnosis of classes will help in determining the rate and attainment which can normally be expected of pupils, and will give normal, understandable, and attainable goals.

In the field of teaching, and particularly in the judgment of teaching, the determination of a sane diagnosis of a class will render the judgment more sound and make the work of the teacher more sure. The teacher who knows the position of his pupils, who knows the point where their further learning must start, who knows what may be expected of them, and who is aware of their fundamental needs, can hardly help dealing with them more intelligently and helpfully.

Determination of potential class ability. The potential ability of a class is the first requisite in the understanding of a class. It is needed as a guide for the teacher in order that he may know what to expect of his pupils as a whole. If a group of pupils has a low ability, the teacher should know it and make allowances; if it is high, he should know it and plan his work accordingly. It is needed also as a protection for the pupils, so that they may not be overstimulated or underestimated. It is also needed for the welfare of the school itself. Comparison of schools and classes on the basis of absolute achievement alone, or the planning of work on such a basis alone, is both unfair and unwise. When potential class ability is known, since this knowledge serves as a guide for the teacher, a protection for the pupils, and an aid in the development of a school, teaching will be improved. The illustration of the combined class used in this chapter is a good example of the values of this type of knowledge.

Determination of potential class ability in the age scale. In the age scale the steps for determining potential class ability are as found below. It will be noted that the steps in the process are such as to make a comparison between the actual potential ability of the class and the potential ability of normal pupils. The same process is characteristic of the technic in other scales.

STEP 1. *Determining the mental ages of the pupils*

The mental ages of the pupils are found through transmuting the raw scores on an intelligence test into mental ages by means of a raw-score-age table as described on page 258.

In Table LXXXVI the mental ages for the pupils shown in Fig. 97, p. 365, are listed.

STEP 2. *Determining the mean of the mental ages*

The mean of the mental ages may be determined either by finding the sum of the mental ages in months and dividing by

the number of ages to be considered, or by the guessed-mean method if a frequency distribution has been made.[1] When there is a small number of cases the former method will probably be found shorter. The addition of the mental ages of Table LXXXVI, divided by 37, the number of ages, results in a mean mental age, as there given, of 148.6 months.

TABLE LXXXVI. MENTAL AGES FOR FIFTH GRADE (DERIVED FROM COLUMN 8 OF FIG. 97)

No.	PUPIL	MA	No.	PUPIL	MA
1	A	123	21	U	162
2	B	147	22	V	136
3	C	158	23	W	138
4	D	143	24	X	169
5	E	118	25	Y	154
6	F	155	26	Z	177
7	G	182	27	AA	128
8	H	162	28	BB	139
9	I	128	29	CC	191
10	J	150	30	DD	143
11	K	166	31	EE	140
12	L	159	32	FF	134
13	M	167	33	GG	126
14	N	127	34	HH	168
15	O	154	35	II	128
16	P	131	36	JJ	132
17	Q	148	37	KK	137
18	R	156			Mean, 148.6
19	S	185			
20	T	138			

STEP 3. *Determining the norm age*

The norm for this class is determined by finding the normal age for the class. This is a fifth grade in May. The average entering age of pupils in this school system is 6–4, and therefore, by consulting Table XXVIII B, p. 232, we find that the normal age at the end of eight months of fifth grade is 139 months. This, then, is the norm age.

[1] See page 182.

STEP 4. *Comparing the norm age and the mean mental age*

The mean mental age is 148.6 months and the norm age is 139 months. This means that this particular class has a mean mental age nearly 10 months greater than the normal age.

Determination of potential class ability in the grade scale. In the grade scale the determination of potential class ability is similar to that of the age scale, with the substitution of grade scores for the age scores.

STEP 1. *Determining the grade scores for the pupils in intelligence* (G_i)

Grade scores in intelligence are determined, as described on page 255, by transposing raw scores on a mental test into G scores. G scores may also be derived from T scores or from a T–G table for an intelligence test.[1] When the G_i scores have been derived for a combined group of sixth and seventh-grade pupils, they would appear as in Table LXXXVII.

TABLE LXXXVII. G_i SCORES FOR A COMBINED SIXTH AND SEVENTH GRADE

SIXTH-GRADE SCORES			SEVENTH-GRADE SCORES		
No.	PUPIL	G_i	No.	PUPIL	G_i
1	A	6.8	16	AA	7.4
2	B	7.8	17	BB	6.4
3	C	6.4	18	CC	4.2
4	D	5.8	19	DD	6.2
5	E	8.2	20	EE	5.4
6	F	9.7	21	FF	5.8
7	G	6.6	22	GG	5.2
8	H	7.8	23	HH	6.6
9	I	5.2	24	II	6.4
10	J	7.0	25	JJ	7.2
11	K	6.2	26	KK	7.0
12	L	8.2	27	LL	8.8
13	M	7.0	28	MM	7.8
14	N	5.4	29	NN	7.0
15	O	7.8	30	OO	7.6
		Mean, 7.06			Mean, 6.60

[1] See page 262, Table XXXVII.

STEP 2. *Determining the mean grade score in intelligence*

The mean grade score in intelligence is found by determining the mean of the various scores for the pupils. In the case cited in Table LXXXVII the mean G_i for the sixth-grade group was found by adding the G_i scores and dividing by 15, the number of cases, and for the seventh-grade group it was found in the same way, the number of cases being equal. The means are to be found at the foot of the columns of G_i scores.

STEP 3. *Determining the grade norm in intelligence*

The grade norm in intelligence is a G score, which is equivalent to the actual grade of the class at the time of giving the intelligence test. In the case cited in Table LXXXVII the test was given in May, and therefore the grade norm in intelligence is 6.8 for the sixth-grade pupils and 7.8 for the seventh-grade pupils.

STEP 4. *Comparing the mean G_i with the G_i norm*

The comparison of the two G_i scores found in the steps above will show the potential ability of the classes in comparison with the grade scores in intelligence of normal pupils in the same grades. The sixth grade has a mean G_i of 7.1 and a norm G_i of 6.8, which indicates that the grade intelligence of the class is equivalent to that of a class three months more advanced. The seventh-grade portion of this class has a mean G_i of 6.6, which is less than that of the sixth-grade portion and one grade and two months below that of normal pupils of the same grade as themselves, since their G_i norm is 7.8. In potential ability the sixth-grade portion of this class is approximately normal, whereas the seventh-grade portion is a year behind normal.

Determination of potential class ability in the T scale. In the T scale the determination of potential class ability is found by a comparison of the mean T scores for the class with the norm T score for classes in similar grades.

STEP. 1. *Determining the T scores for the pupils in intelligence*

These T scores are found by transmuting the raw scores on an intelligence test into T scores as described on page 260. The T_i scores for the pupils in the combined sixth and seventh grades cited previously are as given in Table LXXXVIII.

TABLE LXXXVIII. T_i SCORES FOR A COMBINED SIXTH AND SEVENTH GRADE

SIXTH GRADE			SEVENTH GRADE		
No.	PUPIL	T_i	No.	PUPIL	T_i
1	A	51	16	AA	54
2	B	56	17	BB	49
3	C	49	18	CC	38
4	D	46	19	DD	48
5	E	58	20	EE	44
6	FF	66	21	FF	46
7	GG	50	22	GG	43
8	H	56	23	HH	50
9	I	43	24	II	49
10	J	52	25	JJ	53
11	K	48	26	KK	52
12	L	58	27	LL	61
13	MM	52	28	MM	56
14	N	44	29	NN	52
15	O	56	30	OO	55
		Mean, 52.3			Mean, 50.0

STEP 2. *Determining the mean T_i score*

The mean T_i score can be found easily in this case by adding the scores for each section of the class and dividing the sum for each section by 15, the number of cases in each. This has been done as in Table LXXXVIII, and the mean score there found is 52.3 for the sixth-grade group and 50.0 for the seventh-grade group.

STEP 3. *Determining the T_i norm for the grade*

The T_i norm for the grade is found by consulting a T–G table for the intelligence test which was used to derive the

T scores. The grade level for the sixth grade, since the test was given in May, is 6.8, and for the seventh grade, 7.8. Consulting the T–G table for this test, which is given in Table XXXVII, the T score is 51, for a G score of 6.8, and 56 for a G score of 7.8.

STEP 4. *Comparing the mean T_i score with the T_i norm*

The comparison of the mean T_i score with the T_i norm will show how far above or below the norm the class stands in potential ability. In the case of the sixth-grade class, with a mean T_i score of 52.3 and a T_i norm of 51, the class stands 1.3 T units above the norm for its grade, whereas the seventh grade, with a mean T_i score of 50 and a T_i norm of 56, stands 6 T units below the standard for its grade.

The same result could be obtained through the use of B_i scores, in which case the norm B_i score would be 50, and it would then not be necessary to determine a norm as in the case of the T_i scores. The mean B_i score could merely be compared with 50 to determine the number of T units above or below the norm that the class may be.

Determination of class achievement. The determination of class achievement formed one of the first purposes for the development of standard tests, and remains one of the chief reasons for their continued use. The determination of class achievement enables the adjustment of the pupils' work. It informs the teacher of the extent to which the pupils have progressed and, regardless of the particular standards of the curriculum, shows him the points to which they have attained in their accomplishment. When class achievement in a number of different subjects is considered, the teacher is enabled to make adjustments in the work of the pupils and also to determine the peculiar or general teaching emphases which are needed by the class as a whole. Moreover, such a process is needed as a corrective for the possible errors in the

otherwise unsupported judgment of the teacher with respect to the progress of his class.

Determination of class achievement in the age scale. In the age scale the process for the determination of class achievement consists in finding proper point measures for the scores made by the pupils and comparing these with the norms which have been established.

STEP 1. *Determining the subject ages for the pupils.*

The subject ages for the pupils are found as described on page 212, by reading the ages directly from the raw scores in the raw-score-age table. Such a group of subject ages is shown in Fig. 91 and reproduced here in Table LXXXIX.

TABLE LXXXIX. SUBJECT AGES FOR A FIFTH GRADE

No.	Pupil	Re. A.	Ar. A.	Sp. A.	No.	Pupil	Re. A.	Ar. A.	Sp. A.
1	A	132	130	130	20	T	124	149	135
2	B	144	144	148	21	U	171	161	177
3	C	177	161	171	22	V	119	144	144
4	D	135	153	139	23	W	144	149	139
5	E	128	153	124	24	X	171	171	148
6	F	135	144	116	25	Y	156	171	139
7	G	171	153	159	26	Z	156	144	144
8	H	156	196	154	27	AA	119	144	159
9	I	144	156	124	28	BB	128	144	132
10	J	156	156	135	29	CC	156	146	144
11	K	147	146	142	30	DD	153	176	154
12	L	156	161	156	31	EE	153	124	137
13	M	153	144	156	32	FF	128	165	137
14	N	124	130	139	33	GG	132	161	154
15	O	128	144	124	34	HH	153	165	142
16	P	128	146	126	35	II	128	144	142
17	Q	135	156	148	36	JJ	135	138	130
18	R	156	180	146	37	KK	147	153	146
19	S	201	161	192	Mean . .		145.3	153.0	144.1

STEP 2. *Determining the mean of the subject ages*

The mean of the subject ages may be found by finding the sum of the ages and dividing by the number of ages, or

it may be found from a distribution by the guessed-mean method. By the former method the means are as found at the foot of the columns in Table LXXXIX.

STEP 3. *Determining the age norm for the grade*

The age norm for the grade can be found from Table XXVIII B, according to the method there described, or it may be determined directly from a table of age scores if it has been made up for the particular test. The age norm for this particular group of subjects was determined from Table XXVIII B in the column headed "6–4," and is there found to be 139 months for the eighth month of the fifth grade.

STEP 4. *Comparing the mean subject ages with the age norm for the class*

The direct comparison of the mean subject ages with the norm for the class will show the relation of the achievement of the class with other classes in a similar grade. In this class the comparisons are as follows: the mean reading age is 145.8, or 6.8 months above the norm (approximately half a school year, which by this scale would be 6.5 months); the mean arithmetic age is 153 months, or 14 months above the norm (one month greater than a school year); the mean spelling age is 144.1 months, or 5.1 months above the norm.

Determination of class achievement in the grade scale. When using the grade scale the process is similar to that which has been described above, using grade scores in place of age scores.

STEP 1. *Determining grade scores in subject matter*

Grade scores in subject matter may be found as described on pages 203 and 224 and may be derived either from age scores if the age scale is also being used, or from T scores when the T scale is used. In the following list of grade scores the scores have been derived from T scores and are taken from the same class as that used in Tables LXXXVII and LXXXVIII.

TABLE XC. G SCORES FOR A COMBINED SIXTH AND SEVENTH GRADE

SIXTH GRADE					SEVENTH GRADE				
No.	PUPIL	G_r	G_a	G_s	No.	PUPIL	G_r	G_a	G_s
1	A	5.1	5.8	6.6	1	AA	6.5	11.6	7.7
2	B	8.0	6.8	7.5	2	BB	6.1	9.3	8.8
3	C	4.9	6.8	8.0	3	CC	5.1	6.8	4.9
4	D	7.4	8.6	9.3	4	DD	5.4	6.5	5.8
5	E	6.5	10.0	10.5	5	EE	5.8	8.6	4.3
6	F	4.3	8.0	11.7	6	FF	5.4	5.3	7.0
7	G	7.4	9.3	7.5	7	GG	4.0	7.5	4.7
8	H	8.0	8.6	11.2	8	HH	6.1	6.2	4.3
9	I	13.8	8.6	6.2	9	II	6.5	8.6	4.3
10	J	7.4	6.2	7.7	10	JJ	5.8	7.5	5.8
11	K	5.4	9.3	7.0	11	KK	5.1	8.0	7.7
12	L	8.8	5.8	7.7	12	LL	6.5	10.8	10.5
13	M	7.4	10.8	8.4	13	MM	5.8	8.6	7.5
14	N	6.8	7.2	5.1	14	NN	6.5	7.5	6.8
15	O	6.1	10.0	8.0	15	OO	6.8	10.0	8.8
Mean . . .		7.2	8.1	8.2	Mean . . .		5.8	8.2	6.6

STEP 2. *Determining the mean grade score*

The mean grade score is found, as previously described, either by simple arithmetic or by the guessed-mean method. The columns of Table XC have been added separately for the two classes, and the sums each divided by 15, as there are fifteen cases in each class. The means will be found at the foot of the columns for each class.

STEP 3. *Determining the grade norm*

The grade norm is simply the grade of the class at the time of giving the test. In the case taken in Table XC one class had completed the eighth month of sixth grade, and the other had completed the eighth month of seventh grade. The grade norm for one class, therefore, is 6.8, and for the other, 7.8.

STEP 4. *Comparing the mean grade score with the grade norm*

In the various subjects the comparisons of the mean grade scores and the grade norms is as follows: the seventh grade

has a G_r of 5.8, a G_a of 8.2, and a G_s of 6.6. Comparing these with the grade norm of 7.8, it can be seen that the class is two years behind normal in reading, four months ahead in arithmetic (a difference of two years and four months between reading and arithmetic), and one year and two months below normal in spelling. The sixth grade has a G_r of 7.2, a G_a of 8.1, and a G_s of 8.2. Comparing these scores with the grade norm of 6.8, it can be seen that the class is four months ahead in reading, a year and three months ahead in arithmetic, and a year and four months ahead in spelling. This sixth-grade group, moreover, is above the seventh-grade norms in each field and is ahead of the seventh-grade group in both reading and spelling.

Determination of class achievement in the T scale. In the T scale the determination of class achievement can be made either on the basis of T scores or on the basis of B scores. When T scores are used, separate T norms for the grade must be found; but when B scores are used, the norm is always 50. It is usually more convenient and involves less labor to find the class achievement in terms of T scores unless B scores are to be used for other purposes. It is usually slightly more accurate, since T and G tables for all tests may not exactly coincide.

STEP 1. *Determining T or B scores in the subject matter for all the pupils*

For the fifteen sixth-grade and fifteen seventh-grade pupils in the combined class cited above in Tables LXXXVII–XC, the T and B scores are as shown in Table XCI.

STEP 2. *Determining the mean T scores or mean B scores for the class*

The mean T or B score can be found either from a frequency distribution by the guessed-mean method, or from the sum of the scores divided by their number. In Table XCI the means are given at the foot of the columns.

TABLE XCI. T AND B SCORES IN SUBJECT MATTER FOR A COMBINED SIXTH AND SEVENTH GRADE

SIXTH GRADE

No.	PUPIL	CA	BCR	T_r	B_r	T_a	B_a	T_s	B_s
1	A	11–11	2	43	45	41	43	51	53
2	B	13–0	−2	58	56	48	46	55	53
3	C	11–5	4	42	46	48	52	57	61
4	D	16–0	−12	56	44	56	44	61	49
5	E	12–1	1	51	52	62	63	65	66
6	F	12–4	1	39	40	54	55	69	70
7	G	13–9	−5	56	51	59	54	55	51
8	H	12–1	1	58	59	56	57	67	68
9	I	12–1	1	71	72	56	57	49	50
10	J	12–1	1	56	57	44	45	56	57
11	K	11–11	2	45	47	59	61	53	55
12	L	12–5	1	60	61	41	42	56	57
13	M	12–2	1	56	57	66	67	58	59
14	N	12–7	−1	53	52	50	49	44	43
15	O	12–3	1	49	50	62	63	57	58
		12–5	T Mean	52.9		53.5		56.9	
			B Mean		52.6		53.2		56.6

SEVENTH GRADE

No.	PUPIL	CA	BCR	T_r	B_r	T_a	B_a	T_s	B_s
1	AA	14–1	−6	51	45	70	64	56	50
2	BB	13–6	−4	49	45	59	55	60	56
3	CC	13–8	−4	43	39	48	44	44	40
4	DD	13–11	−5	45	40	46	41	48	43
5	EE	13–9	−5	47	42	56	51	41	36
6	FF	15–3	−10	45	35	38	28	53	43
7	GG	14–9	−8	37	29	51	43	43	35
8	HH	14–1	−6	49	43	44	38	41	35
9	II	15–2	−9	51	42	56	47	41	32
10	JJ	13–11	−5	47	42	51	46	48	43
11	KK	13–9	−5	43	38	54	49	56	51
12	LL	16–1	−12	51	39	66	54	65	53
13	MM	12–11	−2	47	45	56	54	55	53
14	NN	13–8	−4	51	47	51	47	52	48
15	OO	14–1	−6	53	47	62	56	60	54
		14–2	T Mean	47.3		53.9		50.9	
			B Mean		41.2		47.8		44.8

STEP 3. *Determining the T norm*

When using T scores the T norm for the grade in which the pupils are located must be found by consulting the table of T–G norms for each test. In the case given in Table XCI the T norm for the sixth grade is 53 in reading, 48 in arithmetic, and 52 in spelling; for the seventh grade the T norm is 57+ in reading, 52+ in arithmetic, and 56+ in spelling. The B norm is always 50, and it should be remembered that this norm is in terms of age rather than grade. If an age T norm is desired, the T norm for the mean chronological age of the pupil should be found. This mean can be found similarly to the means previously described.

STEP 4. *Comparing the mean T or B scores with the norms*

To compare the mean T scores with the T norm is to indicate, in T units, how far above or below the normal grade T score are the achievements of the class. In the sixth grade the mean T_r is 52.9, the mean T_a is 53.5, and the mean T_s is 56.9; this class is, then, 0.1 T below the norm in reading (a negligible amount), 5.5 T above it in arithmetic, and 4.9 T above it in spelling. In the seventh grade the mean T_r is 47.3, the mean T_a is 53.9, and the mean T_s is 50.9; this class is therefore 9.7T below the norm in reading, 1.9 T above it in arithmetic, and 5.1 T below it in spelling.

It will be noted that these judgments are slightly different from those arrived at in terms of the G scores. The explanation probably lies in the fact that the grade norms of 6.8 and 7.8 used in the G system are inaccurate and that the two classes are not really 6.8 and 7.8 classes. If the school had been reclassified before these tests were given, it is probable that the tests for grades 6.8 and 7.8 would give results more comparable with the age units, as represented by the T or B scores.

In comparing the mean B scores with the norm it should be remembered that the norm in B scores is always 50, so that it is easy to make direct and accurate comparisons. It is also helpful to remember that B scores are in terms of T units throughout, but are in terms of age rather than grade comparisons.

Determination of class effort. As a part of the diagnosis of a class the determination of class effort bears an important relation to the final judgments which can be rendered. It has been shown that the quality of achievement of a class is relative to the ability of the class, rather than in absolute terms of total accomplishment. Therefore, to judge the work of a class properly that work should be judged in relation to the ability with which the class has been endowed to do it. It is also necessary to make a judgment of class effort in order to determine what is wise for a class in terms of future progress. Classes cannot be maintained at a constant level of achievement, and they frequently vary greatly in different types of subject matter. The future progress of a class, particularly with reference to the emphases to be placed upon varying forms of subject matter in future teaching, can be aided greatly by a determination of the class effort. A further advantage in the determination of class effort is that it allows a better judgment of the teacher's work. Teachers are easily misjudged when the total achievement of their pupils is considered solely. If a teacher has a group which is below normal in its achievement, it is a more difficult matter to maintain normal standards than is the case when the pupils are above par. Only too frequently the teacher of pupils who are below par, who maintains his pupils at their optimum learning rate, is criticized in comparison with one whose pupils are above normal and who, although exceeding the standards laid down, is yet maintaining a class rate of learning less than its

optimum rate. The determination of class effort is the proper corrective for this fallacy.

Determination of class effort in the age scale. In the age scale there are several ways in which a satisfactory determination of class effort can be obtained, depending upon the data which are most conveniently at hand.

1. *When achievement ratios have been found in subject matter.* When the teacher has determined achievement ratios in subject matter, the determination of a class effort score is simple, being merely to find the mean of the subject ratios. This has been done in Table XCII, where achievement ratios in reading have been found for each pupil. The mean of these achievement ratios is found at the foot of the table, and the comparison can be made, as in all achievement ratios, with the norm of 100. The reading-score data and the mental ages are taken from Fig. 96.

TABLE XCII. ACHIEVEMENT RATIOS OF FIFTH-GRADE PUPILS IN READING

No.	Pupil	Re. A.	MA	AR	No.	Pupil	Re. A.	MA	AR
1	A	132	123	107	20	T	124	138	89
2	B	144	147	97	21	U	171	162	105
3	C	177	158	112	22	V	119	136	87
4	D	135	143	94	23	W	144	138	104
5	E	128	118	109	24	X	171	169	101
6	F	135	155	87	25	Y	156	154	101
7	G	171	182	94	26	Z	156	177	88
8	H	156	162	104	27	AA	119	128	93
9	I	144	128	112	28	BB	128	139	92
10	J	156	150	104	29	CC	156	191	81
11	K	147	166	88	30	DD	153	143	107
12	L	156	159	97	31	EE	153	140	109
13	M	153	167	109	32	FF	128	134	95
14	N	124	127	97	33	GG	132	126	104
15	O	128	154	83	34	HH	153	168	91
16	P	128	131	97	35	II	128	128	100
17	Q	135	148	91	36	JJ	135	132	102
18	R	156	156	100	37	KK	147	137	107
19	S	201	185	108	Mean . . .		145.3	148.6	98.5

2. *When only subject ages and mental ages are available.* If the subject achievement ratios have not been found, the effort score for a class can be determined directly from the subject ages and the mean mental age, as follows:

STEP 1. *Determining mean subject ages*

For a small class these can be determined by adding the ages for the group of pupils and dividing by the number of pupils. In Table LXXXIX, the reading, arithmetic, and spelling ages for the fifth grade shown in Fig. 96 are listed, with the mean ages at the foot of each column.

STEP 2. *Determining the mean mental age*

This mean score is found as in Step 1 above. In Table LXXXVI the mental ages for the pupils are listed, with the mean at the foot of the column.

STEP 3. *Applying the achievement-ratio formula*

The achievement-ratio formula, as described on page 297, is

$$(100 \times) \frac{\text{Subject age}}{\text{Mental age}} = \text{Achievement Ratio.}$$

When the achievement ratios have been found between the mean mental age and the various subject ages, the results show that the AR is 97 in reading, 103 in arithmetic, and 97 in spelling (secured through comparison of the mean mental age, 148.6, the mean reading age, 145.3, the mean arithmetic age, 153.0, and the mean spelling age, 144.1). It should be remembered that the comparison with the norm is automatic, since the norm for achievement ratios is always 100; therefore ratios below 100 indicate effort less than normal, and ratios above 100 show effort greater than normal.

Determination of class effort in the T scale. When the class effort score is desired and the T scale has been used, there are two ways of making the determination:

1. *When F scores have been determined in subject matter.* If F scores have been determined for the separate units of subject matter (F_r, F_a, and the like, and not F_e or other composite F scores), it is merely necessary to find the mean F score in the subject matter under consideration. This can be done as shown in Chapter XII. It is then merely necessary to remember that normal F scores are 50, and the comparison is automatic. Table XCIII shows a series of T_a, T_i, and F_a scores found for a class in arithmetic, with the mean F_a score shown at the foot of the column.

TABLE XCIII. F_a SCORES FOR CLASS IN ARITHMETIC

No.	Pupil	T_a	T_i	F_a	No.	Pupil	T_a	T_i	F_a
1	A	34	39	45	20	T	44	45	49
2	B	41	49	42	21	U	50	55	45
3	C	50	53	47	22	V	41	44	47
4	D	46	47	49	23	W	44	45	49
5	E	46	38	58	24	X	54	57	47
6	F	41	52	39	25	Y	54	52	52
7	G	46	62	34	26	Z	41	60	31
8	H	70	55	65	27	AA	41	41	50
9	I	48	41	57	28	BB	41	46	45
10	J	48	50	48	29	CC	42	66	26
11	K	42	55	37	30	DD	56	47	59
12	L	50	53	47	31	EE	32	46	36
13	M	41	56	35	32	FF	51	43	58
14	N	34	41	43	33	GG	50	40	60
15	O	41	52	39	34	HH	51	56	45
16	P	42	42	50	35	II	41	41	50
17	Q	48	49	49	36	JJ	38	43	45
18	R	59	52	57	37	KK	46	44	52
19	S	46	63	33	Mean $F_a = 46.5$				

2. *Determination of class effort from B or T scores.* When effort scores have not been found, it is possible to determine class effort scores directly from either T or B scores. The process is the same for both and involves merely the selection of the T or B formula as shown on page 301.

STEP 1. *Determining the mean T_i or B_i for the class*

The T scores in intelligence or the B scores in intelligence for a class should be added and the sum divided by the number of scores, or else a mean should be found from a frequency distribution. The result is a mean T score or B score in intelligence. Table LXXXVIII shows a column of T_i scores for a combined sixth and seventh grade, with the means entered at the foot of the columns.

STEP 2. *Determining the T or B scores in the various phases of subject matter*

The T and B scores in reading, arithmetic, and spelling for the class of sixth-grade and seventh-grade pupils previously cited are as shown in Table XCI.

STEP 3. *Determining mean T or B scores in subject matter.*

These mean scores can be found by addition and division by the number of cases, or by the guessed-mean method. The mean T and B scores for the pupils there considered are shown at the foot of each column in Table XCI.

STEP 4. *Applying the F-score formula to the mean scores*

The F formula (see page 301) is $[T_e \text{ (or } B_e) + 50] - T_i \text{ (or } B_i) = F_e$.

When this formula has been applied to the mean T scores shown in Tables LXXXVIII and XCI, the results are as shown in Table XCIV. The comparison with normal effort is made by remembering that the normal effort score is 50. It should be noted that the F scores derived from T units and those from B units in the same subject matter and with the same pupils are the same, so that both need not be determined for any class.

Summary of diagnosis. Considering the combined sixth and seventh grades from the standpoint of class diagnosis, the sixth-grade group is about normal in ability, whereas the

TABLE XCIV. DETERMINATION OF MEAN F SCORES FOR COMBINED SIXTH-GRADE AND SEVENTH-GRADE PUPILS FROM MEAN T SCORES

	MEAN T SCORES				MEAN F SCORES		
	T_i	T_r	T_a	T_s	F_r	F_a	F_s
For sixth grade	52.3	52.9	53.5	56.9	50.6	51.2	54.6
For seventh grade	50.0	47.3	53.9	50.9	47.3	53.9	50.9

seventh-grade group is nearly a year deficient. The two portions of this class are approximately equal in ability. In achievement the sixth-grade group in this class is above its norm in all three subjects, — reading, arithmetic, and spelling, — whereas the seventh-grade group is two years below its norm in reading, one year below it in spelling, and a little above it in arithmetic. The two groups considered together are not comparable in achievement, and the sixth-grade group is superior to the seventh-grade group throughout. In effort the sixth-grade group is above normal in each subject, whereas the seventh-grade group is well above only in arithmetic, nearly equal to the norm in spelling, and below the norm in reading. It is clear that the seventh-grade group is a slower-moving group than the sixth grade, and if given optimum opportunity the sixth grade would outstrip the seventh. With both groups the teacher could well afford to spend less effort in arithmetic and more in reading. A complete reclassification of the pupils is indicated as a necessity, and the separation of the slower-moving group from the faster-moving one.

Chapter summary. The diagnosis of class achievement is significant in aiding the development of the pupils, in making necessary adjustments in the curriculum, and in influencing the conduct of teaching for the better.

Potential class ability may be determined by using scores in the age, grade, or T scales. In the age scale it is found by

comparing the pupils' mean mental age with the norm age; in the grade scale, by comparison of the mean grade status in intelligence with the norm status in intelligence; in the T scale, through comparing the pupil's mean T score in intelligence with the norm T score in intelligence.

Class achievement may also be determined in the age, grade, or T scale. In the age scale class effort may be found either through determining a mean achievement ratio in subject matter or by finding a mean subject age for the class and applying the achievement-ratio formula with the mean mental age. In the T scale the class effort may be determined either with the T or B scores or with F scores in subject matter. When F scores have been found, the mean F score in subject matter represents the class effort score. When T or B scores are used, it is necessary to determine a mean T_i or B_i score and to find an F score using mean T or B scores in subject matter.

A consideration of the various class ratings in potential ability, achievement, and effort will enable a teacher to adjust his teaching, to judge the quality of the achievement of his pupils, and to make a better estimate as to future emphases and progress. To a supervisor, it will indicate where a teacher needs his constructive help.

SELECTED BIBLIOGRAPHY

BURTON, W. H. Supervision and the Improvement of Teaching, pp. 288–319. D. Appleton and Company, New York, 1922.

BROOKS, S. S. Improving Schools by Standardized Tests, pp. 51–95. Houghton Mifflin Company, Boston, 1922.

MONROE, W. S., DEVOSS, J. C., and KELLY, F. J. Educational Tests and Measurements (Revised and Enlarged Edition), pp. 431–468. Houghton Mifflin Company, Boston, 1924.

PAULU, E. M. Diagnostic Testing and Remedial Teaching. D. C. Heath & Co., Boston, 1924.

CHAPTER XVI

THE DIAGNOSIS OF SCHOOLS

Standard tests and the schools. The technics for determining the relationship of standard testing to the school system have been given in detail in their basic elements in other sections of this book. It is the purpose of this chapter to outline the processes whereby these tests are used to contribute to an understanding of a school system and to aid in the adjustment of its problems. It is not the intention here to detail technics or to elaborate processes, since these processes concern the classroom teacher (for whom this book is prepared) less than they do the administrative and directive forces of the schools themselves. Nevertheless the diagnosis of schools is of interest to classroom teachers, because it is through them that changes, improvements, and corrections must eventually be instituted.

As has been repeatedly shown, there are many elements within the schools which act to influence the character of the school system, and which can be understood better through the use of standard tests. The technics for the diagnosis of pupils and classes and for the classification of pupils, which have been described previously, as well as those for teaching and rating pupils, which are to be treated later, are able to contribute much to the understanding of the work of the schools, and particularly to the adjustment of its various parts.

Influence of variations in pupils and teachers. It has been abundantly shown that pupils differ widely in their reactions to teaching as well as in their native abilities to react. They vary widely from one another, even on the same age levels,

in their achievement in school subject matter as well as in their rate of learning. As pupils vary, so also do teachers. This variation among teachers is much affected by the type and amount of professional preparation which they may have had, by the length of their service in teaching, and the like. These variations result in many complex problems. The effort of the school administration must be to keep the learning and the teaching upon as high a level as possible. To accomplish this it becomes necessary for it to lay down for the guidance of its teachers certain essentials in the form of a curriculum. It then becomes the duty of the teachers to regulate the varying degrees of accomplishment of the pupils so as to approach for all of them as closely as possible to the standards which have been laid down. Such a process, emphasizing constantly the progress of the group, tends toward the maintenance of group standards, which, in turn, may lead to formalism. This process, unless counteracted in some way, although it may appear to have an apparent degree of material success, will lead eventually to a defeat of its own ends by reducing teaching on all levels to a materialistic striving for the acceleration of some who should be retarded, and for the retardation of many who should be accelerated. Such a result brings about both mechanized teaching and unfruitful learning.

In the solution of the problems which are raised in this connection standard tests have become almost indispensable, since it is possible, on the one hand, for group as well as individual attainment standards to be set in terms of many tests in different fields and, on the other, to account for and act as an aid in the adjustment of the variations from those reasonable standards.

One use of standard tests involves the analysis of teaching to reveal its essentials. This can be done through the practice tests which are now available and, to a somewhat less

extent, through the analysis of results from other forms of tests. In the practice tests in particular the elements which make them up are representative of the essentials that teachers should be expected to teach. If, then, the analysis of such tests shows what has not been taught, not only is the corrective suggested, but the remedy is at hand.

The determination of desirable objectives as respects individual cases has been already dealt with. As respects the objectives of the group, the determination has been implied. If a class of pupils has as its objective a given grade level of work as determined by certain standard tests, it has been implied previously that this particular grade level has been set as the standard for the school system for like pupils. In this way standards of attainment in reading, spelling, arithmetic, English composition, handwriting, and the like can be set by a school system for itself in terms of acceptable norms for standard tests and, by using the tests themselves, can determine the extent to which the standards have been met or are being met.

The determination of the needs of pupils upon varying levels of intellect or achievement is a field in which standard testing can play a big and important part. There are few problems in the school system which are more insistent than those of meeting the needs of pupils of varying degrees of intellect and achievement. Standard tests can therefore be used, first, for determining the abilities and rates of learning of the pupils in general. This results in classification of the pupils in so far as they can be properly classified, in terms of both rate of learning and achievement.

There are, secondly, in every school system pupils who by reason of extremely slow rates of learning should not be dealt with in the classes arranged for the mass of the pupils. The work which extremely dull pupils are capable of doing satisfactorily and the means by which such pupils can be taught

differ so much that it has been found expedient for school systems to segregate such pupils for purposes of special instruction and help. If placed in the regular classes of the school system, these pupils tend to become a drag upon the rest of the class in which they are and to diminish the effectiveness of the teaching for all the pupils, including themselves.

Thirdly, pupils with high rates of learning are also found in school systems, and, as revealed through standard tests, these pupils are as numerous as those of lower ability. Such pupils should find the work and rate of teaching in the regular classes of the school system not adapted to their needs, and greater opportunity should be provided for them. Whether this opportunity should be to allow such pupils to progress at their optimum rates in the same fields as normal pupils and thereby complete the prescriptions of the elementary school in a shorter time than normal pupils, or whether it should provide for increased breadth in the curriculum, is still a moot question.[1] This is not a matter which standard tests can settle, but standard tests have revealed the need, and experimentation may lead to the solution.

Influence of new objectives in education. The new objectives which have become evident in education and which were discussed in Chapter III, pp. 18–20, have brought about decided changes in schools and in the conduct of schooling. The emphasis on individualism has been the corrective for the formalism which was discussed in the preceding section. The emphasis on the development of initiative in pupils and on greater freedom of action and of thinking has resulted in an increasing broadening of the curriculum and a reorganization of the course of study.

The results of these changes, as might be expected, have been evidenced by chaotic methods of teaching, because of

[1] For an interesting discussion of this point see B. R. Buckingham's *Research for Teachers* (Silver, Burdett and Company, 1926), pp. 272–274.

the experimentation of teachers with new materials, new ideas, and new aims. As a further result it has been difficult to evaluate and even, in many cases, to identify the goals. The net result has been most clearly evidenced in the condition of the curriculum. Changes in time devoted to various subjects, the greater freedom allowed to teachers in their teaching, changed emphases even within well-defined units of subject matter, such as reading and the like, have tended to produce great differences in the curriculum, even within a single school system.

Standard tests are now being used with effectiveness in the solution of such difficulties. In many branches of study which were formerly given a greater proportion of time than is possible today, such as reading, arithmetic, and spelling, standard tests can be used for the evaluating of various methods and for the measurement of progress. The use of such tests in these fields is aiding in the removal of the superfluous elements of teaching, is placing proper emphasis on the elements of true importance, and is making possible a better adjustment of time on the part of both teacher and pupils. The use of standard tests, with the wealth of knowledge which has developed as a result of their use, has transformed the teaching of spelling. It has made possible definite goals for pupils to attain at different stages in their progress through school and has supplied teachers with materials of instruction better and more complete than they have had. What is true in the field of spelling is equally true in arithmetic, handwriting, reading, and like school skills, and is becoming increasingly true in additional fields.

Influence of the public interest in schools. The past few years have seen the increasing development of an interest in the schools on the part of the public. The occasion is the mounting total cost of education, as compared with former years, and the consequent more careful scrutiny of the pur-

poses for which the moneys are spent. In all public activities the same condition is to be seen — an increasing tendency to spend the public moneys with wisdom.

The burden of proof for the efficiency of the educational process lies with the administrators of the public schools. There is little tendency to curtail the expenditure of moneys for school purposes that are clearly valuable, but there is legitimate question raised as to the expenditure of such moneys for ill-defined or unproved purposes.

Standard tests can be used effectively to show what the schools are doing for the moneys they receive. Through the use of such tests it is possible to point out (see pages 443–448) where teaching is being done efficiently and where there are difficulties, and it is possible to provide correctives for the difficulties and thereby promote efficiency and prevent waste.

Probably one of the more effective means of proving the efficiency of the schools is through the educational publicity which the use of standard tests makes possible. The public, interested as it is in the public schools, is nevertheless only too often ignorant of what they are doing and of the worth of their efforts. Publicity, of a character which will at one time show what the schools are doing and point out the needs with which they are confronted, can do much both to aid the schools and to prove their worth.

A third service of standard tests in promoting economy in the schools lies in the adjustment of time and the uses to which the elements of the curriculum are put. When it can be shown that given amounts of time in one school produce as good results as greater amounts of time in another (as is frequently the case even when the average abilities of the pupils concerned are equal), the opportunity is given to the school administration to discover the causes and to make such readjustments as to increase the efficiency of the school which has been devoting the larger amount of time. Such a change

would in its turn release time which might better be devoted to another and more worthy purpose.

A fourth use of standard tests in this connection lies in the possibility which they give for the increase in coördination of work as between schools within the system. With the variations in types of teaching and of school population which are characteristic of city schools, and with the increasing proportion of pupils who graduate from the elementary schools to gather in junior and senior high schools, differences as among schools become rapidly apparent. The use of standard tests, first, to measure these differences and, secondly, to aid in equalizing them makes it possible for pupils to transfer from school to school and from elementary schools to the higher schools with greater probability of success and without the necessity for repetition of work or for the loss of essential elements in development.

Influence of social and mental differences among pupils. Among the changes which have acted upon the schools, some of the more important have been of social character. In many school systems, owing to the aggregation of foreign-born peoples in well-defined districts, it is becoming more usual to find in certain schools high proportions of children of foreign-born parentage. These pupils, with their differences in home tradition and of racial background from those of pupils of native-born parents, are not only handicapped in their learning but need many emphases and aids not so necessary for the native pupil. In these instances as much school time as possible should be devoted to developing the necessary social and racial adjustments. To make the teaching time spent with the pupils as effective as possible in the basic elements of the curriculum, standard tests can be used effectively and can provide a check on the development of the pupils.

A second form of social adjustment which is necessary among school pupils is due to the increasing leveling of pro-

vincial barriers. The development of the moving picture, the automobile, and the radio has had several results with school pupils. Populations drift more widely than in the past and know better the activities of people in more remote sections of the country. To take account of these increased contacts of both parents and children and to make for greater equality of opportunity for pupils, it is necessary that school systems should have more in common. This can be accomplished through the medium of the standards set up by standard tests, whereby the work of one school system, the rates of progress established for the pupils, the objective results of teaching, and the like can be compared with the results in other communities or with other sections of the country. Such a process, when it involves a leveling up of standards rather than a leveling down, is healthful for the schools and desirable for the communities.

A third form of social need is that involved in providing opportunity for pupils of varying intellectual or mental capacities. The emphasis that is now being placed on the relation of the individual to the social group of which he is a part takes cognizance of the varying abilities of pupils and their varying needs. This phase of the use of standard tests has been dealt with in terms of the measurement of the effort of pupils, and is particularly significant from the point of view of the school system in the possibilities for the segregation of certain types of pupils. When a pupil accomplishes as much as he could reasonably be expected to accomplish in the accepted forms of school study, and, even then, lacks a development which meets the minimal standards of the community, it is evident that he needs special attention. Standard tests can, on the one hand, be used for the location of such pupils in the schools, and can, on the other hand, become an aid in measuring the extent of the development which results from the segregation.

A fourth social need, closely related to all the foregoing discussions of this chapter, is one related to the teachers themselves. There is probably no key position in the school system ultimately of greater significance than that of the classroom teacher. There is, then, probably no more important work of the school system than the constant stimulation and encouragement of teachers for the improvement of their work. All that part of supervision not actually connected with the inspection of schools and the maintenance of the minimal standards of instruction is concerned with the development of the teachers in service. As a guide for the needs of such teachers, as an evidence of where they can profitably spend time and effort in further study, and as a means for bringing to them the significance of increased development, standard tests offer to the administration of school systems a means which is impersonal and which is eagerly sought by most teachers. Since these ratings of teachers are indirect, since they lack the bias of personal prejudice, since they reveal the actual needs of pupils as well as of teachers, and, particularly, since they make possible comparisons with the results obtained by other teachers elsewhere under similar conditions, they are acceptable to teachers as evidences of their own success or failure. The study which teachers must make of their pupils, the review which they must make of their methods of teaching, the judgments which they must make of their decisions in the classroom, all result in increased efforts and in more effective teaching.

The school survey and standard tests. The diagnosis of schools has been developed in organized fashion through the development of the school survey. In a general school survey it is the practice of a group of experts in educational procedure to investigate the conditions in a given community as respects the schools and to make recommendations for the improvement of the situation. Such surveys are becoming

more numerous and have proved their worth in terms of the advantages which they bring to the schools, the teachers, and the public.

The school survey contemplates a measurement of the entire school system as it affects the administration of schools, the development of school-housing, the development of the curriculum, and the treatment of pupils. To illustrate the method and use of standard tests in the development of a school survey, and particularly the relation of the facts discovered and recommendations made to the interests and needs of the classroom teacher, the following sections quote characteristic portions taken from a typical and excellent survey. Only those portions of the survey are utilized which are the result of the giving or analysis of some standard tests given to the school pupils, or which bear directly on the problem.

Standard tests and the pupils. In this survey standard tests in many different fields were given to the pupils, and the test results were analyzed in combination with other classroom and school conditions for the determination of various aspects of the educational process. It is unnecessary to show the entire process, but a few quotations will show the method used and the results from the standpoint of the pupil.

1. *The sectioning of pupils.* The report shows from a study of the placement of the pupils in the elementary grades that there was a high proportion of over-ageness in the schools, which was considered to be the result of "improper promotion policies, the lack of attention to age of entrance, and to inadequate adjustment of the curriculum to the abilities of boys and girls."[1] In terms of each of these criticisms the investi-

[1] Report of the Survey of the Schools of Duval County, Florida (including the city of Jacksonville), made by the Institute of Educational Research, Division of Field Studies, Teachers College, Columbia University, George D. Strayer, Director, p. 169. Bureau of Publications, Teachers College, Columbia University, New York City, 1927.

gation made is cited, and recommendations are made for improvement of conditions by correcting the promotion policy so as to reduce the necessity for so much repetition of teaching, by attempting to bring pupils into the schools at the optimum time, and by adjusting the curriculum to better advantage for the pupils. In one section of this latter recommendation the following is a characteristically helpful discussion:

A study was made of three sections of one sixth-grade class which had been sectioned upon the basis of teachers' judgments and previous grades. Table 46 shows shifts that would result if this same group were resectioned on the basis of a good measure of learning rate.[1]

TABLE 46. SHIFTING THAT WOULD RESULT IF PUPILS WERE GROUPED UPON MORE ADEQUATE MEASURES OF LEARNING RATE THAN THE JUDGMENT OF THE TEACHERS

Present Fast Section
60% should remain in section.
40% should be shifted to the average section.
None should be shifted to the slow section.
Present Average Section
27% should remain in average section.
33% should be shifted to the fast section.
40% should be shifted to the slow section.
Present Slow Section
60% should remain in the slow section.
33% should be shifted to the average section.
7% should be shifted to the fast section.

In the fast and slow sections a little more than half the children were properly grouped; in the other section only about one fourth were properly grouped. It is recommended that the groups be selected upon the basis of probable learning rate as determined from standard intelligence and educational tests. Even the best

[1] The measure of learning rate used by the survey staff is an average of the intelligence quotient and the educational quotient (in this book this measure is called the promotion quotient; see page 285).

teachers at times misjudge some of the boys and girls. Table 46 shows that 7 per cent of the slowest group had the ability to do the work of the brightest section.

Even when the grouping is done on the basis of standard tests, there will be times when a pupil should be shifted up or down, and the system must provide for the necessary arrangements which will see that such shifts are made. A lack of flexibility is one of the commonest shortcomings of ability sectioning. Teachers of the slower sections often hesitate to lose their better pupils through change to a brighter section, and teachers of the brighter sections often feel that they do not want to impose an added burden upon the load of teachers who are struggling with the slower pupils. Added to all this is the fear that the teacher may be mistaken about the pupil's ability after all. The record of the amount of resectioning that occurred in the elementary schools during the first semester of this school year indicates that in Jacksonville, as elsewhere, the difficulties above outlined have resulted in too little resectioning. There is only one elementary school in which any amount of resectioning occurred. It is recommended that the organization for pupil classification provide for such reclassification periodically.[1]

2. *Special types of pupils.* In addition to the class problems considered above, there is further consideration in this survey of the needs of special types of pupils. The first quotation below indicates the recommendation with respect to mentally retarded pupils; the second, to a bright but retarded pupil.

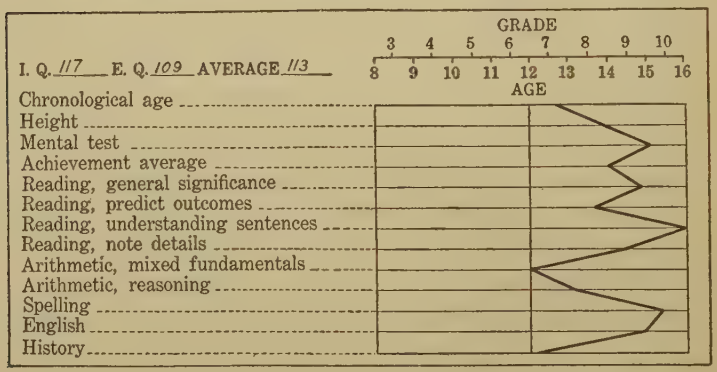
One of the most significant movements in education during the last few years is the segregation of pupils who, all experience shows, are unable to profit by ordinary school work. The development of intelligence tests has made it possible to discover these pupils at an early age. In common with many other communities, Jacksonville has taken a very significant step in this direction. There are now 129 pupils in classes for mentally retarded pupils, and significant work has been done toward giving them school

[1] Duval County Survey, pp. 195–196.

opportunities by which they can profit each year of their school life. The experience of cities which have worked along this line longer than Jacksonville shows, however, that what Jacksonville has done is only a good beginning. It is safe to say that there are at least 350 boys and girls in the schools of Duval County who would profit by such special class instruction. The survey staff recommends that the excellent beginning be extended until this service reaches all the boys and girls who could profit by it.

CHART 20

A Girl who should be given a Special Promotion[1]



There is need for further work in the development of the opportunities offered those who are already enrolled in special classes. Teachers of these classes should be encouraged and assisted in obtaining more training for this highly specialized type of service. More adequate equipment is likewise desirable. The pupils should always be given pleasant rooms. The policy of having these classes in the larger schools is an excellent one. It is recommended, however, that the older boys and girls be housed in the junior high schools where broader opportunities in hand work may be given them.[2]

[1] This is an able sixth-grade pupil who is a year over age for her grade and is achieving on the average of two years ahead of her grade. Broken line shows this pupil's record. Vertical line represents the grade standard.

[2] Duval County Survey, pp. 167–168.

Chart 20 gives the educational record of a 6A pupil who, due to repeating and late entrance to the first grade, is now two years behind her proper grade. The test records show that in the majority of her tests she is more than a year ahead of her present grade standard. Doubtless acceleration in this case would not only place this pupil nearer to her age group, but would serve as an encouragement to make up lost time.[1]

Standard tests and the curriculum. Many of the data found in a school survey report which results from the analysis of standard tests given to the pupils have a direct bearing upon the curriculum that is taught. They reveal the rates of learning of the pupils, the standards maintained in the various units of subject matter in the different grades, and, if carried far enough in interpretation, they indicate the adjustments that might be made. The survey from which we have quoted considers these questions in detail for arithmetic, reading, language, and the like. A part of the discussion concerning reading will show the extent of the analysis and its value to the classroom teacher.

The Thorndike-McCall Reading Test was given in all city schools and three rural schools. Table 25 shows that except in Grade 6B, where the achievement is approximately two months behind standard, the work in reading is, on the whole, better than standard. The same table, however, shows that while this is true for the situation as a whole, it is not true for the country schools, but that the latter are seven months behind in the fourth grade and a year and a half behind in the sixth grade.

Table 24 shows the time in minutes per week given by a Jacksonville school and by other schools in various parts of the country to various school subjects. It will be seen that, whereas in Grade 3 Jacksonville gives slightly less than average time to reading, the proportional time given to reading in Jacksonville gradually increases so that in Grades 5 and 6 it is considerably greater than in other cities. This would lead one to expect the reading achievement

[1] Duval County Survey, pp. 192–193.

TABLE 24. COMPARISON OF NUMBER OF MINUTES ALLOTTED PER WEEK TO VARIOUS SUBJECTS AND ACTIVITIES IN THE SECOND HALF OF GRADES 3 TO 6 IN A TYPICAL ELEMENTARY SCHOOL FOR WHITE CHILDREN IN JACKSONVILLE, FLORIDA, WITH THE MEDIANS FOR 93 CITIES IN THE UNITED STATES,[1] 1927

SUBJECTS	GRADE 3 HIGH			GRADE 4 HIGH			GRADE 5 HIGH			GRADE 6 HIGH		
	Jacksonville	Median	Number of Cities	Jacksonville	Median	Number of Cities	Jacksonville	Median	Number of Cities	Jacksonville	Median	Number of Cities
Language and Grammar	105	145	90	200	154	91	225	167	92	225	173	89
Arithmetic	200	207	91	225	219	91	225	220	92	225	220	89
Geography	105	76	68	225	138	85	225	156	86	225	157	84
Reading	275	303	90	225	226	90	225	201	88	225	169	85
Art and Drawing	75	89	87	75	87	87	75	90	87	75	89	86
Spelling	75	103	89	75	101	89	75	100	90	75	82	86
Penmanship	75	79	90	75	79	90	75	78	91	75	78	85
History	105	61	40	100	89	58	175	104	73	150	118	70
Citizenship and Civics		34	25		39	31		37	36		40	32
Social Science		77	12		108	11		130	12		140	12
Phonics	50	59	41	25	58	14		55	6		28	5
Literature		59	29		65	27		70	28		73	26
Nature Study		39	40		43	36		49	33		45	25
Handwork		58	21		62	17		61	13		63	11
Manual Arts	105	64	9		63	10		70	27	90	73	40
Music	75	80	90	75	93	90	75	91	91	75	88	88
Physical Training	75	103	78	75	102	80	75	100	83	75	100	78
Health Education		48	66		52	69		52	72	150	53	72
Household Arts		65	1		63	6		67	22	90	76	31
Supervised Study		107	12		130	14		130	15		140	14
Projects and Activities		38	3		35	3		28	3		25	2
Recess	125	127	77	125	123	76	125	109	72	75	108	69
Opening Exercises	50	55	77	30	54	74	30	54	76	30	55	67
Unassigned Free Time	75	60	36	95	62	37	20	56	36		56	35
Miscellaneous	105	50	25	50	50	20	50	39	19	55	37	21

[1] The median for each subject was calculated on the basis of the number of cities reporting time allotted to that subject in the course of study. The number of cities in each case appears in the table next to the median. For example, the median time spent on manual arts in Grade 3 is 64 minutes. This is based on 9 cities which allot time to manual arts. Eighty-four cities report no time to this subject in Grade 3. The 93 cities considered are cities having from 25,000 to 100,000 population. Forty are in the Eastern section of the United States, 25 in the Great Lakes section, 11 in the Southern section, 10 in the Great Plains section, and 7 in the Western section.

of the sixth grades in Jacksonville to be considerably farther above the standard than in the fourth grade. This, it has been shown, is not the case; instead, in the upper grades, Jacksonville spends more time on reading than other cities, and accomplishes less.[1]

TABLE 25. RESULTS OF TESTS GIVEN IN 11 ELEMENTARY, 3 RURAL ELEMENTARY, AND 2 JUNIOR HIGH SCHOOLS, DUVAL COUNTY, FLORIDA, MARCH, 1927

NAME OF TEST	GRADE	NUMBER CASES	LOW SCORE	HIGH SCORE	DUVAL COUNTY 25 PERCENTILE	DUVAL COUNTY 75 PERCENTILE	JACKSONVILLE MEDIAN SCORE	DUVAL COUNTY RURAL MEDIAN SCORE	DUVAL COUNTY TOTAL MEDIAN SCORE	GRADE STANDARD
Thorndike-McCall Reading	4B	142	1.5	13.3	3.88	5.53	4.66		4.67	4.15
	4A	239	1.5	15.0	4.28	6.06	5.25	3.93	5.08	4.65
	6B	120	2.0	13.3	5.17	6.79	5.98		5.97	6.15
	6A	251	3.0	14.4	5.57	7.85	6.89	5.11	6.72	6.65
Morrison-McCall Spelling	4B	142	1.5	8.4	3.35	4.57	3.96	3.55	3.90	4.15
	4A	228	2.0	10.5	3.84	5.16	4.47	3.75	4.40	4.65
	6B	136	3.0	11.2	5.02	6.81	5.91	5.67	5.86	6.15
	6A	226	2.0	12.5	5.56	8.11	7.01	7.05	7.02	6.65
	8B	38	3.5	12.5	6.44	8.06	7.5		7.5	8.15
	8A	83	4.5	13.0	6.90	9.80	8.19		8.19	8.65
Woody-McCall Arithmetic Mixed Fundamentals	4B	170	2.0	6.4	3.69	4.87	4.39	4.00	4.36	4.15
	4A	207	2.0	9.3	4.23	5.24	4.80	4.31	4.76	4.65
	6B	134	4.0	10.0	5.48	7.16	6.42	5.5	6.36	6.15
	6A	236	3.0	12.6	5.73	7.68	6.72	5.63	6.63	6.65

This discussion of the general levels of reading within the scope of the survey is followed by a description in greater detail of the results of a further group of tests, four forms of the Gates Silent Reading Tests. A portion of this discussion is as follows, omitting the table upon which the discussion is based:

It will be seen, that Jacksonville pupils are up to standard in each of the grades tested other than 6B Grade in Reading to Understand Precise Directions; that is, in ability to read with

[1] Duval County Survey, pp. 137–139.

exactness and precision, to select accurately relevant details while subordinating other facts however interesting, and to retain without distortion the precise directions to be executed.

Two other forms of the Gates Reading Tests are Reading to Appreciate the General Significance of a Paragraph and Reading to Predict the Outcome of Given Events. The latter type of reading tested requires, in addition to a grasp of the general significance of the passage, also a special type of analysis of the facts in order to predict which of several possible events will be most likely to happen next. In both of these tests the upper-half grades are in every instance practically or quite up to standard, and the B sections are below standard from two to nine months.

In Reading to Note Details, which requires not a grasp of some one general characteristic or one subtle implication of the whole passage, but an independent treatment of three different details or aspects of the passage, Jacksonville pupils in each of the three grades tested fall far below standard. The deficiency ranges from three months in Grade 4 A to more than three years in Grade 8 B. Evidently Jacksonville pupils in general do not need so much emphasis on reading of the type necessary to understand precise directions, but they do very decidedly need training in ability to note details as they read.[1]

Standard tests and the teachers. It should be perfectly evident from the foregoing discussion that the results of a survey of this type are of extreme importance in many phases of the teacher's work. Even such results as have been quoted cannot fail to be of real help to the interested teacher, but there are other ways as well in which he can receive aid and stimulation. One of these is in the opportunities that are given to the supervisors and principals to consider and discuss with their teachers the problems which have been raised. One of the problems is to create that atmosphere of mutual trust and understanding, based upon some mutual recognition of needs, which will result in helpful discussion. Standard tests given wisely to pupils supply such a basis for discussion; and

[1] Duval County Survey, pp. 139–140.

the means may be in private conference, which for some types of difficulties (such as those pertaining only to a single classroom) is desirable, or in open teachers' meetings, grade meetings, building meetings, and the like.

Commenting upon this phase of the situation, the survey report states as follows:

> Along with the knowledge of individual differences and, to a great extent, supplying such knowledge, has developed the movement for scientific measurement of educational conditions and products. A large part of the school procedure is still based upon tradition and opinion, but notable progress has been made in subjecting educational questions and theories to the test of scientifically determined fact. Educators know some of the laws of learning and the relative merit of certain methods of teaching with as much certainty as the physician knows the effects of tested medicines and diets. It is possible to measure the intelligence of children and their abilities in certain types of school work with a high degree of exactness.
>
> This is a new field in which new discoveries and new instruments are announced with bewildering rapidity. Much of the published material is in a form too technical to be understood by most teachers or principals. It is no wonder that the results of research affect school practice slowly and in spots rather than in a systematic way. Nevertheless, a school system which is awake must in some way keep up with educational progress. It cannot be content to use methods which have been proved to be ineffective nor to neglect instruments which permit exact measurement of school work to replace belief and assertion.
>
> Individual members of the staff will, on their own initiative, make some use of the announcements of research workers. They should be encouraged to do so. Nevertheless, to enable the school system to take full advantage of progress in educational measurement, the services of specialists are needed. There should be on the staff persons who have the requisite ability and training to interpret the results of research and make it available to teachers in a practical form. The supervisors who come in direct contact with teachers must make themselves familiar with educational

measurement and coöperate with teachers in applying it to classroom procedure. They will need, however, the assistance of specialists in research to evaluate the published material and to put it in form which teachers can use.[1]

Standard tests and the curriculum. In treating of the adjustment, improvement, and extension of the curriculum, quotations were made with respect to pupils,[2] which show the method of deriving the results from the testing program, and particularly the means whereby points for special emphasis as to rate of learning, quality of material included in the curriculum, standards, and the like can be determined and emphasized.

It is, of course, definitely understood that revision of a curriculum does not rest upon standard tests: it merely can begin there. To determine what correctives should be applied in teaching to eliminate the difficulties revealed by standard tests necessitates the change of materials and the revision of methods of work on the part of both teachers and pupils. It may involve a change in attitude toward the work. The following quotation from the survey report previously cited shows how the staff has interpreted the situation with respect to spelling in terms of both materials and method. It is to be understood that the changes resulting should be measured at some time later to prove their worth.

> The scientific studies in the teaching of spelling and the selection of word lists have greatly reduced the amount of time necessary for the teaching of spelling. What words to teach and the most economical method of teaching are now known.
>
> The studies by Thorndike, Ashbaugh, Jones, and Graham have shown the need for individual work and the waste incurred if all children are permitted to spend the same amount of time upon a given list of words.

[1] Duval County Survey, pp. 43–44. [2] See page 443.

Most of the teaching methods observed in spelling were commendable, but the choice of words is questioned. The selection of proper materials for the spelling work should be taken care of immediately, since today there are available several scientifically built lists of words with proper grade placement.

Dr. Horn gives as the five most important points of method:

1. The teacher must test her pupils on each lesson before they begin to study.

2. Each pupil should study only the words which he misspelled on this test.

3. Each pupil must be taught an economical method of studying.

4. Each pupil must see clearly what progress he is making.

5. Controlled reviews must be definitely distributed to prevent the lapse of habits after they have been initiated.[1]

If these principles are to be applied, more time must be spent in teaching the child how to study his spelling lessons. Directions should be worked out for each child such as:

1. Look at the word and pronounce it.

2. Look at each part as you pronounce it.

3. Say the letters.

4. Close your eyes and see if you can remember how it looked. Check to see if correct.

5. Look at the word, close your book, and write the word. Check. If a mistake has been made, write the word correctly three times.

The time given to testing should be relatively short. Each child should know his score, keep a record of misspelled words, and review frequently.

Some teachers find it advantageous to make spelling work a group problem, taking any well-established list of words and selecting the ones most pertinent to the needs of their grades. The children plan the work, care for individual drill, and ask the teacher's coöperation only in the giving of tests and in scoring.[2]

[1] E. Horn and J. Ashbaugh, *New Horn-Ashbaugh Speller*, p. 12. J. B. Lippincott Company, 1926.

[2] Duval County Survey, pp. 255–256.

Enough has been said to show the importance of standard tests in the survey of schools, and particularly of the contribution made by the use of such tests in the understanding of school conditions. The relationships of these tests to the pupils, the teachers, and the curriculum could be extended to include their value to the schools as a whole in terms of the location and definition of needs in school opportunities, particularly in terms of pupils with physical disabilities. Studies of the efficiency of various schools for the improvement of conditions for pupils transferred from one school to another, and studies of pupils for classification, for promotion, or for grading purposes, and the like, could be expanded, but the ways in which such results are obtained have been outlined not only in many of the quotations here given but also in other chapters of this book.

A section could also properly be devoted here to the relation of standard tests to the public through the medium of the school survey. In general the results of tests may be safely given to the public if in so doing care is taken to indicate the interpretations which should be made; but in no case is it desirable that the results of tests of particular pupils be given to the parents or to the pupils so that invidious comparison will result. This is particularly true of the results of intelligence tests. The results of achievement tests may be used by pupils, and undoubtedly the parents might be advised in intelligible terms of the results as well, but rather in terms of grade status than in terms of ages and the like. It is salutary, therefore, for the public to be informed through a survey of this type of its position in the matter of its schools. No general survey is undertaken unless the public is desirous of improving its schools, and the public is therefore entitled to the facts. Reliable publicity (which, of course, means intelligible interpretation of the technical facts) is the surest guaranty of a public appreciation of the efforts for economy

in education, the professional improvement of teachers, and efficiency in learning by the pupils.

Chapter summary. Various factors combine to influence the character of the school system, and many of these can be determined and measured through the medium of standard tests.

Variations in pupils and teachers can be measured and the problems raised in the school system can be solved through the setting up of adequate individual as well as group attainment standards, and through the promotion of ways for meeting the needs of pupils with varying degrees of intellect and attainment.

The changes in the objectives in education have placed on the school system not only the burden of attempting to meet the needs of these new objectives but also that of justifying the experimental methods which are being used. The influence of public interest in the schools, which has increased greatly within the past few years, places an additional necessity upon the schools not only to see that children are educated but also to see that the moneys spent for that education are expended as wisely as possible.

In addition to these factors the social factors in the community are also becoming more influential in determining the course of the school development. The presence of children of foreign parentage, of pupils with wider ranges of interest, and of better-informed classroom teachers makes the problem of the development of the curriculum and the treatment of pupils different from what it formerly was. Standard tests can be used to aid in the solution of these problems.

An effective means of illustrating the broad use and great value of the intensive use of standard tests in a school system is illustrated in the description of a school survey, wherein the relation of the tests to the pupils and teachers, the curriculum, the public, and the schools is shown.

SELECTED BIBLIOGRAPHY

BROOKS, S. S. Improving Schools by Standardized Tests. Houghton Mifflin Company, Boston, 1922.

MCCALL, W. A. How to Measure in Education, pp. 149–168. The Macmillan Company, New York, 1922.

MONROE, W. S., DEVOSS, J. C., and KELLY, F. J. Educational Tests and Measurements (Revised and Enlarged Edition), pp. 431–468. Houghton Mifflin Company, Boston, 1924.

MORT, P. R. The Individual Pupil. American Book Company, New York, 1928.

OTIS, A. S. Statistical Method in Educational Measurement, pp. 267–288. World Book Company, Yonkers-on-Hudson, 1925.

PAULU, E. M. Diagnostic Testing and Remedial Teaching, pp. 96–127. D. C. Heath & Co., Boston, 1924.

TRABUE, M. R. Measuring Results in Education, pp. 144–171. American Book Company, New York, 1924.

WILSON, G. M., and HOKE, K. J. How to Measure (Revised Edition), pp. 356–387. The Macmillan Company, New York, 1928.

CHAPTER XVII

THE RATING OF PUPILS

Difficulties in the rating of pupils. Much has been written with respect to the difficulties, injustices, and unfair practices concerned with the rating of pupils. There has been little controversy as to whether pupils should be rated at all, although many reasons for rating have been given; the main question has been rather how to rate them.

In substance the main difficulties with the rating of pupils by teachers may be summarized under three main ideas. Difficulties have arisen, first, through the variability in the judgment of pupils and the products of pupils by their teachers; in the second place, injustice has been caused by variability in the standards used for rating; in the third place, unfairness has been caused through variability, or rather invariability, in what was expected of pupils.

There has been much proof offered to show the variability and fallibility of the teacher's judgment. Given two equally good products by different pupils, teachers have judged them as of different worth; given the same product, two teachers have judged it to be of different worth. Such criticisms are now of common knowledge, and form one of the main reasons for the general acceptance and faith on the part of teachers in standard tests. Teachers know themselves to be fallible and have seized upon standard tests to aid them in resolving this difficulty.

The variability in rating standards (particularly as it results in an injustice in the rating of pupils) is less obvious than the variability in judgment, and has not received such a con-

centration of attention. So long as the teacher is dealing with a single class, — a homogeneous unit in appearance, — and so long as he is dealing with a rating scale which seems to have fairly general acceptance both by himself and the individuals upon whom it is used, the rating scale seems to be satisfactory and invariable. It is thus that a teacher uses a rating scale in large measure. It is only when extremes in the use of the scale are cited — instances rarely met with by the average teacher — that the injustice in the scale becomes apparent and the injustice is shown to be variable merely in degree and not in kind. If an essay type of test in history is given to an eighth-grade group of pupils, and the pupils are rated by the teacher who taught them, neither the teacher nor the pupils would be inclined to protest the results very greatly; if, on the other hand, a group of fourth-grade pupils were to be given the same test and rated by the same teacher on the same basis, the protest would be great, and would be justified. If the eighth-grade group received A's and B's, the fourth-grade group would, on the same basis of grading, receive, at most, D's, with little more to hope for.

As between classes in the same school, and even as between classes taught by the same teacher, other things being equal, the same injustice is done; and the rating standard — the meaning of A, or of "passing," of or "80 per cent" — varies from teacher to teacher, from class to class, and from subject to subject.

Until recently variability in achievements of pupils has received even less attention than have the two difficulties previously discussed. Enough has been said both here and elsewhere regarding the differences of individuals in ability and the relation of ability and achievement in terms of effort. It should be obvious that such a condition should in some way be taken into account in rating pupils, but it rarely has been. Ordinarily class ratings on either of the two more com-

mon bases of rating — the system of grades, such as A, B, C and the like, or the popular "percentile" system — are accepted by pupils as well as by teachers as evidences of relative excellence of accomplishment. In school after school, in grade after grade and (most unfortunately for the pupils) in home after home, excellence of accomplishment is measured in terms of excellence of grade. The pupil who "gets" A is uniformly praised, and the pupil who receives E or F or D, or whatever lies at the low end of the scale, is uniformly blamed. There is no question that, as generally allotted, these grades are given in terms of the relative excellence of the work of the pupils within any given class. This fact has been established both by tradition and by edict. But the fact that an A may not mean a high degree of personal success for a pupil and that D or C may mean much higher relative accomplishment for another is rarely questioned, as it should be.

These three difficulties in rating can be helped of solution, if not completely solved, by considering the bases upon which the difficulties are founded and by taking steps to remedy them. Variability in the judgments of the teacher can be corrected through taking means to make those judgments as highly objective as possible. Variability in standards can be corrected through attempting to find a fixed standard. Variability in the achievements of pupils can be corrected through determining what should be expected of pupils. In all these ways standard tests can aid in the resolution of some of the difficulties which have been summarized.

Relation of standard tests to rating of pupils. Standard tests, since they have grown in part directly from the objections to the rating of pupils which were raised, in certain respects are designed to eliminate those objections.

The variability in the judgment of the teacher is almost completely eliminated in standard tests. As the technic of making such tests develops, the tests are becoming more and

more objective in form, in content, in scoring, and in their results. Whenever there is universality of material, there is the opportunity for the construction of standard tests which can eliminate almost wholly the subjective errors of the classroom teacher; where there is not universality in content, in skill, or in means, then tests are relatively less standard and also less objective.

In the matter of variability of standards, standard tests have as yet little to offer, save in narrow fields. They are leading toward absolute standards. They have approached absolute standards in arithmetic, reading, writing, and a few related fields; but in history, geography, science, and the like absolute standards are as yet unapproached. However, even with absolute standards, or with the scales of relative standards such as have been described in earlier chapters, a solution of the rating system for pupils in classes is not found. Possibly some time in the future both teachers and pupils will be satisfied with such standards; but the time is not yet, since although they are useful and accurate they are not universally understood. It is even doubtful if in some cases they should be understood by pupils.

In terms of variability in achievements, standard tests, largely through the intelligence tests, have furnished the first practical means whereby proper programs of learning can be planned. The entire effort-score technics, whether they are the result of classroom tests or standard tests, are directed toward this end and are used increasingly by teachers to determine proper standards of learning for individuals.

As has been said, standard tests, so far as the tests and technics previously described are concerned, are limited in their use as rating elements for pupils' achievements in two important ways: on the one hand, they do not cover all phases of school work, although the field they do cover is wide and is becoming increasingly wider; on the other hand,

although they can set satisfactory age, grade, and individual standards, they do not furnish suitable rating standards for the judgment of pupils' work in terms of accepted and traditional ratings.

There is great question as to whether it is even desirable that the traditionally accepted ratings of pupils should be continued at all. There is much in connection with the percentile and letter-grade rating systems which is undesirable and even pernicious. They can lead to great abuse and misunderstanding. Nevertheless they are almost universally used and as universally accepted by teachers, parents, and pupils, as well as required by administrators. If a satisfactory means could be found whereby the results of standard tests could be transmuted into terms of present grading units, the values of the great objectivity of these tests would be extended to help toward relieving one of the onerous duties of the teacher. The transmutation in any satisfactory form must depend upon the establishment of a fixed standard rather than a variable rating standard.

Establishment of a fixed rating standard. Any rating standard outside the pupil, unless it be an absolute standard, is variable. As has been shown, if the standard of rating is placed within the test, it will vary with the ease or the difficulty of the test. Thus, when a teacher gives a so-called "easy" test, the grades are likely to be high; when the test is "difficult," the grades are likely to be low. It has also been said that if the standard is placed in terms of the opinion of the teacher, the standard is also variable. This is what occurs when it is stated by a teacher that his subject (this occurs more frequently in high school or college) is more difficult than another teacher's. In this case the standard is variable in terms of the teacher's opinion. Another variable standard is undoubtedly the variability as among teachers, one of whom may be considered to "mark hard" and another to

"mark easy." Parents are well aware that their children recognize these differences among the teachers. No matter what basis may be selected, if it lies outside the pupil and is not absolute, it is variable.

The only fixed standard of rating for any group of pupils consists in the status of that group of pupils. This merely means that pupils should be judged by what they have done, and not by what a teacher might think they were taught or might wish they had learned. Such standards as these latter are variable, but throughout all the vicissitudes of learning, teaching, subject matter, and schooling the pupils remain the same. Those who are subjected to superior teaching are fortunate; those who are not are merely unfortunate, not more dull, and are not to be penalized for the mistakes of others. The intellectual status remains the same; the average of the class best represents the class. Tests may vary and opinions and subject matter may also vary, but the pupils are the same, and it is through themselves that they should be judged. The rating of pupils finds its standards within the pupils, not outside them, and the pupils themselves find their ratings in terms of other pupils in the same group.

It is convenient to make distinctions as between pupils. These relative distinctions within the group, like the more inclusive fixed standard itself, must find their meanings in terms of some fixed standard which can become a series of proportions of the group itself. "A" can have only two possible true and invariable meanings: it can have an absolute meaning, such that any pupil who achieves a given rating automatically is deserving of an A; or it can have a relative meaning, such that any pupil who achieves a status within a given proportion of a class deserves an A. The former is practically impossible; the latter is feasible and desirable.

With a fixed standard once determined as the status of the group itself, and with the fixed meanings for the distinc-

tions as among the ratings once determined, the essentials of a fair-rating technic are established and there remains only the determination of the technic itself.

The technic must on the one hand take account of the fixed standard and the fixed discrimination, and on the other hand allow for the character of the group. Strictly speaking, the group of pupils does not change; but one group of pupils has many contacts with another group, and the two groups may exhibit different characteristics. So far as may be possible, allowance should be made for these varying characteristics of groups, and fairness should be exhibited in dealing with them.

A technic for transmuting standard test scores into letter grades. The technic which is described below aims (1) to fix a definite standard in terms of the status of the class, (2) to fix a proportional rating scale in terms of a normal class, and (3) to modify the scale in terms of the variability of an actual group from the normal.

STEP 1. *Determining a letter-grade system*

To select a letter-grade system some decision should be made as to what is the character of the group with which it is to be used. If the group of pupils in ability forms a frequency surface which is similar to the normal curve (see page 190), one of the systems given below which is based upon the normal curve should be selected; however, if the group is in ability a minus-skewed group (see page 191), wherein there are more pupils of higher ability than of lower ability (that is, more pupils above the mid-point of the base line than below it), then one of the systems which is based on minus-skewed curves should be selected. When once the type of group to be graded has been determined, as can be done from an inspection of a frequency surface of the raw scores or T scores of an intelligence test, any convenient letter-grade system may be selected.

TABLE XCV. PERCENTAGE GROUPINGS. LETTER-GRADE SYSTEMS 1 THROUGH 8 (BASED ON NORMAL CURVES)

LETTER GRADE	SYSTEMS							
	1	2	3	4	5	6	7	8
A	3	4	5	6	7	8	10	15
B	23	24	24	24	24	24	24	22
C	48	44	42	40	38	36	32	26
D	23	24	24	24	24	24	24	22
ED	3	4	5	6	7	8	10	15

The letter-grade systems, as given in Table XCV, were derived in the following manner. The base line of a normal curve was divided into five equal parts, based upon plus and minus three P.E.[1] (the P.E. (probable error) is a measure similar to the S.D.), so that for System 1 the upper fifth of the base line, when produced to the curve, cut off exactly 3 per cent of the curve; for System 2 the upper fifth of the base line cut off 4 per cent of the curve; and so on. The percentages of the curve included within each of the remaining fifths of the base line were then calculated for each of the systems, with the percentages resulting as given in Table XCV. The result is a series of systems of percentages of discriminatory grades which are based on equal amounts of deviation from the mean of the groups and are also based on the normal distribution. These systems, therefore, are so derived as to make the selection of any one of them provide for a fixed standard for any normal group of pupils as well as

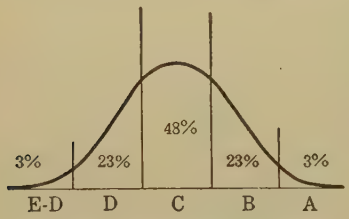


FIG. 102. Derivation of Letter-Grade System No. 1

[1] See M. R. Trabue, *Measuring Results in Education*, pp. 449 ff. American Book Company, New York.

fixed discriminatory proportions of letter grades. Fig. 102 shows graphically how System 1 was derived.

The letter-grade systems shown in Table XCVI were derived from various forms of minus-skewed curves, where the proportion of A grades corresponded to those of systems 1–8, but where the remaining proportions varied with the type of curve used. These proportions are derived empirically and are given because many teachers desire to use systems based upon minus-skewed curves as being more fair for their pupils. This is frequently true, and particularly so for the upper grades of the elementary school wherever pupils have not been accelerated through reclassification.

TABLE XCVI. PERCENTAGE GROUPINGS. LETTER-GRADE SYSTEMS 9 THROUGH 16 (BASED ON MINUS-SKEWED CURVES)

LETTER GRADE	9	10	11	12	13	14	15	16
A	3	4	5	6	7	8	10	15
B	22	23	24	25	26	29	31	32
C	54	52	50	48	46	42	38	32
D	20	19	18	17	16	15	14	13
ED	1	2	3	4	5	6	7	8

In making a selection of a system for use with a group of pupils, any one of the sixteen systems here given may be chosen, or any other system can be made, provided only that the teacher is as consistent as possible in making grades according to the system selected. Only in this way can grades A, B, and the like acquire real meaning. One of the systems given in either Table XCV or Table XCVI should be found applicable in most cases.

STEP 2. *Listing the raw scores, T scores, or ages in rank order*

When the scores of the pupils on the test have been derived, either as age scores, raw scores, or T scores (not

quotients or B scores), the next step in the transmutation to letter grades is to list the scores in order, from best to poorest, as described on page 146.

A list of raw scores in spelling for a group of pupils, arranged in order, is shown in Table XCVII.

TABLE XCVII. RAW SCORES IN SPELLING FOR FIFTH GRADE ARRANGED IN RANK ORDER

PUPIL	SCORE	PUPIL	SCORE
S	47	HH	30
U	43	II	30
C	41	D	29
G	37	N	29
AA	37	W	29
L	36	Y	29
M	36	EE	28
H	35	FF	28
DD	35	J	27
GG	35	T	27
B	33	BB	26
Q	33	A	25
X	33	JJ	25
R	32	P	23
KK	32	E	22
V	31	I	22
Z	31	O	22
CC	31	F	18
K	30		

STEP 3. *Calculating the number of scores in each letter-grade group*

In Step 1 a letter-grade system was to be selected from either Table XCV or Table XCVI. This system gives the percentages of scores in each letter-grade group. For any particular set of pupils, therefore, the next step in the transmutation of scores to letter grades is to determine the number of scores for each group. This is done by multiplying the number of individuals in the class by each of the percentages given in the system selected in Step 1.

The group of pupils for whom this list of spelling scores was derived formed a slightly minus-skewed group. It was thought best, then, to use one of the letter-grade systems based on the minus-skewed curve, and System 13, shown in Table XCVIII, was selected as being desirable.

TABLE XCVIII. LETTER-GRADE SYSTEM 13

LETTER GRADE	PER CENT
A	7
B	26
C	46
D	16
E	5

Each of the percentages in Table XCVIII was multiplied by 37, the number of scores for the pupils shown in Table XCVII, with the result as shown in the third column of Table XCIX; 7 per cent of 37, the proportion of the A ratings, is 2.59; 26 per cent of 37, the proportion of the B ratings, is 9.62; and so on.

STEP 4. *Making a preliminary adjustment of the proportions*

It will be noted in Table XCIX that the results of the calculations in Step 3 are not in whole numbers. The proportion of A ratings, according to the system, should be 2.59. This is impossible to transmute, as the scores to be rated are whole numbers. An adjustment of the theoretically desirable ratings must therefore be made, to bring them to whole numbers. The teacher may make this adjustment in any way that he sees fit, keeping as closely to the desirable ratings as possible. In this case the proportion of 2.59 A ratings is changed to 3, of 9.62 B ratings to 9, of 17.02 C ratings to 17, of 5.92 D ratings to 6, and of 1.85 E ratings to 2. The result is shown in Table C.

TABLE XCIX. PROPORTION OF RATINGS FOR 37 PUPILS

LETTER GRADE	PER CENT	PROPORTION
A	7	2.59
B	26	9.62
C	46	17.02
D	16	5.92
E	5	1.85

TABLE C. FIRST ADJUSTMENT OF RATINGS TO DISCRETE UNITS

LETTER GRADES	PER CENT	PROPORTION	ADJUSTMENT
A	7	2.59	3
B	26	9.62	9
C	46	17.02	17
D	16	5.92	6
E	5	1.85	2

STEP 5. *Applying the ratings to the scores*

The number of ratings in each letter-grade division can now be applied to the scores, and the scores be transmuted to the new ratings. The highest three scores, of 47, 43, and 41, for Pupils S, U, and C, respectively, would be rated as A, the nine scores next lower would be rated as B, the seventeen below them as C, the next six as D, and the final two as E.

TABLE CI. APPLICATION OF FIRST ADJUSTMENT TO SPELLING SCORES OF TABLE XCVII

PUPIL	SCORE	LETTER GRADE	PUPIL	SCORE	LETTER GRADE
S	47	A	HH	30	C
U	43	A	II	30	C
C	41	A	D	29	C
			N	29	C
G	37	B	W	29	C
AA	37	B	Y	29	C
L	36	B	EE	28	C
M	36	B	FF	28	C
H	35	B	J	27	C
DD	35	B	T	27	C
GG	35	B			
B	33	B	BB	26	D
Q	33	B	A	25	D
			JJ	25	D
X	33	C	P	23	D
R	32	C	E	22	D
KK	32	C	I	22	D
V	31	C			
Z	31	C	O	22	E
CC	31	C	F	18	E
K	30	C			

STEP 6. *Making the final adjustment of the ratings*

It will be noted that by the method given above, Pupil X, with a raw score of 33, is included with the C ratings, and that Pupils B and Q, with the same scores, are grouped with the B ratings. Also, Pupil O, with a score of 22, is grouped in the E ratings, whereas Pupils E and I, with the same score, are grouped with the D ratings.

An adjustment should be made in these cases, the obvious one being to group the score of Pupil X among the B ratings and the score of Pupil O among the D ratings, as shown in Table CII.

TABLE CII. SECOND ADJUSTMENT OF SPELLING SCORES OF TABLE XCVII

PUPIL	SCORE	LETTER GRADE	PUPIL	SCORE	LETTER GRADE
S	47	A	HH	30	C
U	43	A	II	30	C
C	41	A	D	29	C
			N	29	C
G	37	B	W	29	C
AA	37	B	Y	29	C
L	36	B	EE	28	C
M	36	B	FF	28	C
H	35	B	J	27	C
DD	35	B	T	27	C
GG	35	B			
B	33	B	BB	26	D
Q	33	B	A	25	D
X	33	B	JJ	25	D
			P	23	D
R	32	C	E	22	D
KK	32	C	I	22	D
V	31	C	O	22	D
Z	31	C			
CC	31	C	F	18	E
K	30	C			

Determination of failure. It should be emphatically stated that no mechanical scheme for rating pupils should include within it the automatic classification of a pupil as a failure

in any particular form of work. Unfortunately the grading systems traditionally in use have come to have meanings attached to their various ratings which are not truly in accord with any except certain, and too often neglected, facts. A rating scale such as one of those which were described on page 460 does not necessarily mean that the rating of A means "Superior" or "Excellent" or that the rating of B is "Fair" or "Good" and the like. Any study of the ratings of pupils in comparison with their abilities to learn, by means of effort scores, will show that frequently the pupils with the higher letter-grade rankings are those with the lower effort rankings, showing that for them the letter-grade rankings are deceptive. Moreover, failure is as relative as the rankings themselves, and bears a relation not so much to the group of pupils as to the work that has been done.

It is conceivable that an entire class could fail to complete what was required of them; it is also conceivable that an entire class could be successful in exceeding what was required of them. An E rating or an F rating, however, has come to mean not the lower proportion of the class in relative achievement (the only real meaning it should have), but rather it has come to connote failure. Under these circumstances (since the meaning of failure is so firmly attached to the rating of E), in using a letter-grade system of this kind it is wise to consider the E grouping as the group where failure is most likely to occur and not as a group in which everyone is a failure. It is necessary, then, for the teacher to scrutinize with particular care the work of the pupil who receives an E rating for the purpose of determining whether his work should be classified as "failing" or as satisfactory. If the pupil has completed the minimal requirements for pupils of his grade, his work should from the grade standpoint be considered satisfactory; if he has not done so, it should be considered unsatisfactory. Thus, if after such an examination of the

pupils' record it is decided that he has not completed the minimal essentials, the grade of E should be allowed to stand as such evidence against him; if, on the other hand, it is judged that the minimal essentials have been satisfactorily met, then, for the sake of the pupil, the letter-grade rating should be changed to D. Only in this way can fairness be established through the medium of this rating system.

It is possible for a teacher to derive help from the grade scores of pupils in standard tests to determine the status of the pupil. For example, in this spelling test the grade scores[1] range as low as 3.5 (for the score of 18). When such grade scores have been found, and if pupils are classified in their proper grade levels, then there should not be a wider variation than five months below the grade norms in a class such as this. This particular class has a grade norm of 5.8, which should range from possibly 5.3 to 6.3 at this time of year (May). If, then, a pupil has a grade score in spelling below 5.3, his work, regardless of whatever other letter grade may be given him, should be considered as below the standard for that grade. As this grade is constituted at this time (when the test was given and before reclassification), there were several pupils below this rating. After reclassification, there would be only one pupil (pupil O) who would be classified in a grade where his spelling score would be below the proper grade level.

Transmutation of scores into percentile ratings. Little has been said with respect to the traditionally used percentile scale for the rating of pupils. Little should be said, since this scale is rapidly passing out of use, particularly in the elementary schools. It is subject to all the misinterpretations and objections that hold for the letter-grade system, and it has one additional great disadvantage: it is not only based upon a variable standard, such as the difficulty of the test or the

[1] See column 16 of Fig. 94, p. 350.

opinion of the teacher, but it is confused with being a measure of absolute attainment. Thus, not only have certain portions of the percentile scale become associated with "Excellent," "Good," and "Poor" attainment, but also a mythical point, varying in different places, has become fixed as a "passing" point. This may be variously placed at 40 per cent, 60 per cent, 70 per cent, 75 per cent, and the like, and the remainder of the scale is adjusted to it.

Regardless of the difficulties and obvious disadvantages of this scale, it is still in use in places where standard tests are also used, and some teachers (usually by school-board regulation and rarely by desire) must report their findings with respect to pupils in such terms.

To transmute raw scores from standard tests into percentile ratings, the attempt must first be made to make the percentile scale conform to a fixed standard, which, as in the case of the letter-grade scale, should be the status of the class with which it is to be used. The technic for making the transmutation is similar to that described for transmuting raw scores into grade scores and the like.

STEP 1. *Defining the meanings of the groupings of the percentile scale in terms of letter-grade groupings*

The meanings of the parts of the percentile scale above the "passing" point are usually similar to those usually attached to the letter-grade scale; thus, a characteristic scale may be shown as in Table CIII.

STEP 2. *Selecting a suitable letter-grade system as a base for the percentile groupings*

The selection of a letter-grade system should be made as described on page 459. Let it be supposed that the same letter-grade system as previously selected has been chosen in this case, and that the same group of scores is to be transmuted into percentile scores.

STEP 3. *Determining the number of scores in each letter-grade and percentile grouping*

This is done as in the preceding sections, and the final grouping of scores would be as shown in Table CII, p. 465, adjusted to meet the conditions of the test. Table CIV shows the score ranges for the percentile groups.

TABLE CIII. A PERCENTILE-RATING SYSTEM

LETTER GRADE	PERCENTILE SCALE
A	95–100
B	85–94
C	75–84
D	70–74
E	Below 70

STEP 4. *Constructing a transmutation sheet*

A transmutation sheet should be constructed similarly to that described on page 204. The lower horizontal edge of the sheet should be labeled with the various percentages, as shown in Fig. 103, and the vertical edge of the sheet should be labeled with the various scores. Then the score points corresponding with the percentile groupings should be placed upon the sheet, as shown in Fig. 103, and the line drawn as there shown.

TABLE CIV. SCORE RANGES FOR PERCENTILE GROUPINGS

(From Table CII)

LETTER GRADE	PERCENTILE RANGE	RAW-SCORE RANGE
A	95–100	41–47
B	85–94	33–40
C	75–84	27–32
D	70–74	22–26
E	Below 70	Below 22

STEP 5. *Reading directly the percentile ratings*

The percentile ratings for each raw score can then be read directly from the transmutation sheet. Where a horizontal raw-score line crosses the intersection of the graph line and a percentile vertical line, the percentage for that score is the percentage represented by the vertical line. The completed table of percentages for these scores, read to the nearest whole number, is given in Table CV.

The value in percentage points of the score in the E grouping, which has been omitted, should be found by the same process as was described in an earlier section. If the work of that pupil is below the standard for his grade, the rating should be made below the "passing" point; if his work is

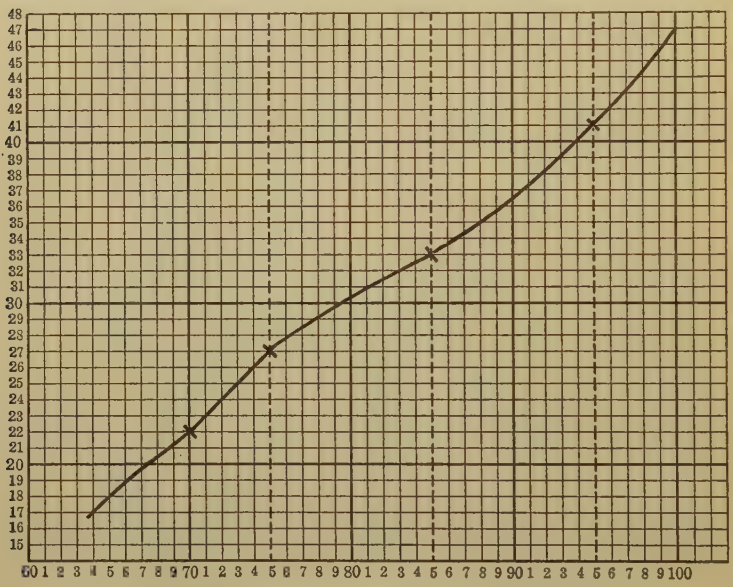


FIG. 103. Transmutation Sheet for Percentile Scores

satisfactory, then, regardless of the system used, the grade should be made at least as good as the "passing" point.

Chapter summary. Three difficulties are found in the traditional forms of rating pupils by means of examinations in terms of letter-grade or percentile scales. The first of these is the variability caused by the subjective judgment of the teacher in the scoring of papers and the results of the learning. This difficulty is almost completely resolved through the use of standard tests, whereby the judgments of the

TABLE CV. COMPLETED PERCENTAGE RATINGS

PUPIL	RAW SCORE	PERCENTAGE RATING	PUPIL	RAW SCORE	PERCENTAGE RATING
S	47	100	HH	30	80
U	43	97	II	30	80
C	41	95	D	29	78
			N	29	78
G	37	90	W	29	78
AA	37	90	Y	29	78
L	36	89	EE	28	76
M	36	89	FF	28	76
H	35	88	J	27	75
DD	35	88	T	27	75
GG	35	88			
B	33	85	BB	26	74
Q	33	85	A	25	73
X	33	85	JJ	25	73
			P	23	71
R	32	83	E	22	70
KK	32	83	I	22	70
V	31	81	O	22	70
Z	31	81			
CC	31	81	F	18	65 or 70
K	30	80			

teacher are reduced to a minimum and the tests themselves are made inclusive. A second difficulty is that of the variability of the standards by which pupils are graded. If the standard used for grading lies outside the pupil, it varies with the basis used and thereby becomes unfair to the pupil. This difficulty is resolved if the standard selected is that which is set up by the pupils themselves and which varies exactly and only as they vary, thus becoming fair to them. A third difficulty is that of the variability of achievements of pupils. This difficulty can only be solved by measuring the achievements of pupils in terms of their abilities to achieve. This necessitates a rating of pupils in addition to the traditional relative rating, and is found in terms of the effort scores which have already been described.

The technic for transmuting raw or other scores from standard tests into letter-grade ratings is given in detail. It consists in determining a letter-grade system as a fixed standard, adjusting the proportions of each rating thus found to discrete units, applying these adjusted ratings to the scores to be rated, and, finally, readjusting the final ratings in the few points where injustices due to the mechanical adaptation may arise.

The technic for transmuting the scores into percentile ratings involves the determination of the fixed standard and the determination of the adjusted limit of the scores and percentile groups. When these have been determined, a transmutation sheet can be made and the percentage ratings read directly.

In both forms of pupil ratings the question of failure, as it is ordinarily known, must be adjusted, not on the basis of a mechanical system of ratings, but rather in terms of the success with which a pupil may have met the minimal requirements of the work. The only use of the lower portion of the rating scale labeled "E" is to reveal the location of possible failures. It is not intended to create automatic failures. If it is decided that a pupil who at first receives an E rating has met the minimal requirements of the work, this rating should be changed to one which would not be confused with failure. The determination of failure is not an attribute of a rating scale.

The rating of pupils by means of standard tests should be considered as merely supplementary to that which can be derived from classroom tests. Although the standard-test ratings are undoubtedly more accurate than those derived from classroom tests, as has been shown, the forms of standard tests are less numerous than would be sufficient for inclusive school ratings.

SELECTED BIBLIOGRAPHY

BUCKINGHAM, B. R. Research for Teachers, pp. 295–329. Silver, Burdett and Company, New York, 1926.

CUBBERLEY, E. P. The Principal and his School, chaps. xix, xxiv. Houghton Mifflin Company, Boston, 1923.

MCCALL, W. A. How to Experiment in Education, pp. 45 ff. The Macmillan Company, New York, 1923.

ODELL, C. W. Traditional Examinations and New-Type Tests, pp. 141–174. The Century Co., New York, 1928.

OTIS, A. S. Statistical Method in Educational Measurement, pp. 267–288. World Book Company, Yonkers-on-Hudson, 1928.

RUGG, H. O. "Teachers' Marks and Marking Systems," *Journal of Educational Administration and Supervision*, February, 1915.

RUSSELL, C. Classroom Tests, pp. 283–328. Ginn and Company, Boston, 1926.

TRABUE, M. R. Measuring Results in Education, pp. 480–485. American Book Company, New York, 1924.

WILSON, G. M., and HOKE, K. J. How to Measure, pp. 551–565. The Macmillan Company, New York, 1928.

CHAPTER XVIII

TESTING IN TEACHING

Relation of testing and teaching. Testing and teaching have always been intimately associated, but until the introduction of more refined methods of school measurement, testing followed rather than preceded teaching. It is now possible for both standard and classroom testing to lead to teaching and not merely to result from it. It becomes, therefore, a new tool for the teacher and one which can be of great use.

Among other uses, it is possible through standard tests to determine for any class or any group of pupils reasonable objectives in certain forms of study and to present these objectives to pupils as goals for achievement. It is also possible, by means of practice tests in certain fields, to keep a constant check on the development of the pupils, to give help to needy ones at a time when it is most needed, and to aid pupils in making their own determinations of their success. By means of classroom tests a teacher can supplement his use of standard tests, can give pupils a means of judging their own products and progress, and can determine as often as necessary wherein pupils have developed. By means of either standard or classroom tests it is possible for a teacher to make divisions of his pupils for purposes of more efficient or more effective instruction, and thereby to make provision for the more uniform and consistent development of all the pupils.

Determination of reasonable objectives, by using the age scale. As has been implied in earlier chapters, the development of pupils is dependent upon many factors other than

mere innate ability. In addition to native ability, there are such factors at work as home environment, prior teaching, rate of work, habits of study, and the like. Of these, one of the greatest influences is that of prior learning. The extent to which a pupil has learned marks the beginning of his new learnings, and therefore predicates, in great measure, what his future can be. Taken into consideration with his ability to learn (see page 248), the knowledge of the extent of prior learning makes possible a reasonable assumption as to the future progress. Therefore, since it is possible to find in objective terms a close approximation of the pupil's true ability and his normal rate of learning as well, it is possible to use these measures as an aid in determining reasonable objectives for his future efforts.

In the age scale the technic of accomplishing this end is easily acquired by teachers. It requires merely a sure knowledge of the use of the measures and a field where there are at least two equivalent forms of tests.

Let it be supposed that a teacher has a number of pupils whose reading ability is variable and who need to be encouraged in their development of that ability. Let it also be assumed that this teacher wishes to have these pupils progress at their optimum rate in reading during the ten months of the school year that they are to be under his direct supervision. The first consideration would be the selection of a suitable reading test. As was implied above, this test should have two or more equivalent forms, because the plan is to give one test at the beginning of the learning period, to calculate the amount of progress which should be made and what accomplishment the pupils should make by the end of the period, and to check the progress at the end of the term decided upon by means of the second test. More than two forms of a test are desirable, so that a teacher can check the progress of the pupils at more frequent intervals if necessary.

The technic which follows is carried out with three pupils, but can easily be carried out for the entire class. It is assumed that the teacher will use such data as are of importance in helping the pupils to understand the objectives which each is expected to attain.

STEP 1. *Determining the chronological age of the pupils at the beginning of the period*

The chronological ages of the pupils at the beginning of the learning period should be determined in months, as outlined on page 94. Table CVI shows in the first column the names of the three pupils under consideration, and in the second column their chronological ages expressed in months.

TABLE CVI. TECHNIC FOR DETERMINING OBJECTIVES IN READING (SHOWING PUPILS AND THEIR CHRONOLOGICAL AGES)

PUPIL	CA
A	142
B	137
C	116

TABLE CVII. TECHNIC FOR DETERMINING OBJECTIVES IN READING (ADDITION OF INITIAL READING SCORES)

PUPIL	CA	*IReS*
A	142	*17*
B	137	*19*
C	116	*18*

STEP 2. *Determining the initial reading scores* (*IReS*)

The initial reading scores are determined by giving one form of a reading test and scoring it for each pupil. In this case Form 1 of the Thorndike-McCall Reading Scale was used, and the three pupils achieved the raw scores as shown in the third column of Table CVII.

STEP 3. *Determining the initial reading age* (*IReA*)

The pupils' initial reading age is found by transmuting the raw scores on the reading test into their respective reading ages. In the Thorndike-McCall Reading Scale this can be done by means of a table of ages which is to be found in

the manual of directions published for the test. If a test has been used where such a table was not available, it could easily have been made according to the directions for making the raw-score-age table which is given on page 212. When the reading ages for the three pupils under consideration were found, they were placed in the fourth column, as shown in Table CVIII.

STEP 4. *Determining the initial mental ages*

The initial mental ages are necessary in order to find achievement ratios and, possibly, intelligence quotients. It should be the purpose of the teacher to stimulate the pupils to read, but not to overstimulate them. The mental ages will be necessary data for this purpose.

TABLE CVIII. TECHNIC FOR DETERMINING OBJECTIVES IN READING (ADDITION OF INITIAL READING AGES)

PUPIL	CA	IReS	*IReA*
A	142	17	*132*
B	137	19	*139*
C	116	18	*135*

The initial mental age may be found either from an intelligence test, if it has not already been found, or it may be found, as described on page 263, if the intelligence quotients are known. When these mental ages have been found they should be placed in the fifth column, as is shown, for the three pupils cited, in Table CIX.

TABLE CIX. TECHNIC FOR DETERMINING OBJECTIVES IN READING (ADDITION OF MENTAL AGES)

PUPIL	CA	IReS	IReA	*IMA*
A	142	17	132	*127*
B	137	19	139	*137*
C	116	18	135	*148*

STEP 5. *Determining the initial achievement ratio (IAR)*

The initial achievement ratio is the effort score for the pupils at the time the reading objectives are established. It

is an indication for the teacher of the extent to which a pupil has been accomplishing as he might, and will aid the teacher in the judgment necessary in determining the objectives. He should try to have the pupils approach more nearly to achievement ratios of 100, their optimum expressions of effort.

The achievement ratio, found as described on page 295 by dividing the subject age by the mental age, has been found for each of the three pupils, as shown in column 6 of Table CX.

TABLE CX. TECHNIC FOR DETERMINING OBJECTIVES IN READING (ADDITION OF ACHIEVEMENT RATIOS)

PUPIL	CA	IReS	IReA	IMA	*IAR*
A	142	17	132	127	*104*
B	137	19	139	137	*101*
C	116	18	135	148	*91*

STEP 6. *Determining the intelligence quotients*

The intelligence quotients may be found directly from an intelligence test,[1] or may be calculated from the mental ages of a test previously given if the date of the test is known.[2] When the intelligence quotients for the three pupils were found, they were entered in the seventh column, as shown in Table CXI.

TABLE CXI. TECHNIC FOR DETERMINING OBJECTIVES IN READING (ADDITION OF INTELLIGENCE QUOTIENTS)

PUPIL	CA	IReS	IReA	IMA	IAR	*IQ*
A	142	17	132	127	104	*89*
B	137	19	139	137	101	*100*
C	116	18	135	148	91	*128*

[1] See page 269.

[2] See page 263.

STEP 7. *Estimating the final reading age (Est.FReA)*

It is now necessary to make some determination as to where the pupils should be at the end of the period selected for the attainment of the reading objectives. Other things being equal, it is reasonable to suppose that a pupil should progress at his potential rate of learning. However, his previous rate of learning may be either greater or less than his potential rate of learning. Therefore factors are known to have entered to make his actual rate of learning differ from his potential rate. The objective of the teacher should be to stimulate a rate of learning such as would bring the pupil more nearly to normal. The estimated final reading age will neglect these extraneous factors, and the final achievement ratio when found will show the teacher the extent of his efficiency in overcoming them.

The final reading age, which can be estimated and set up as an objective for any pupil, may be found from the following formula, which is based on the potential rate of learning of each pupil and the length of the term decided upon for the attainment of the objective.

$$\frac{(IQ \times \text{length of term})}{100} + \text{initial age} = \text{estimated final age.}$$

For Pupil A the substitution becomes

$$\frac{(89\ (IQ) \times 10\ (\text{months}))}{100} + 132\ \ (IReA) = 141\ (Est.FReA).$$

It will be noted that the intelligence quotient has been divided by 100, because it was originally multiplied by 100 to remove the decimal point.[1] The estimated final reading ages for the remaining two pupils have been calculated by the same formula and have been placed in the eighth column of Table CXII.

[1] See page 270.

TABLE CXII. TECHNIC FOR DETERMINING OBJECTIVES IN READING (ADDITION OF ESTIMATED FINAL READING AGES)

PUPIL	CA	IReS	IReA	IMA	IAR	IQ	*Est. FReA*
A	142	17	132	127	104	89	*141*
B	137	19	139	137	101	100	*149*
C	116	18	135	148	91	128	*148*

STEP 8. *Estimating the final reading score (Est. FReS)*

The final reading score, as a more understandable objective than the estimated final reading age, may be estimated by reversing the process of finding the reading age. It is merely necessary to find the estimated reading age on the table of raw scores and ages used in Step 3, and there locate the raw score in reading for that age. In the table of ages given with the Thorndike-McCall Reading Scales the raw scores and ages are not complete enough for as close a determination as one would like, and therefore it is wise to select the nearest reading score for the estimated age. This difference will be noted in the case of Pupil B in comparing the actual final reading age and estimated final reading age with the respective reading scores.

When the estimated final reading scores have been found they can be entered in the ninth column of the table, as shown in Table CXIII.

It should be noted here that if the term for the attainment of the objectives is not an entire school year, that portion of the formula should be adapted to suit the conditions. If seven months, for example, is the term to be used, that number should be used in the formula in place of 10. In the same way, if the teacher wishes to determine an estimated score or age at any portion of the originally planned term, it can be determined by the same formula, using the changed amount of time.

TABLE CXIII. TECHNIC FOR DETERMINING OBJECTIVES IN READING (ADDITION OF ESTIMATED FINAL READING SCORES)

PUPIL	CA	IReS	IReA	IMA	IAR	IQ	EST. FReA	*EST. FReS*
A	142	17	132	127	104	89	141	*21*
B	137	19	139	137	101	100	149	*23*
C	116	18	135	148	91	128	148	*23*

STEP 9. *Determining the final reading score (FReS)*

All these portions of the table may be determined, and should be, at the beginning of the teaching process. They, and especially the estimated final reading score, should be utilized to determine the objectives and degrees of effort necessary to attain them for each pupil, and should be utilized as a guide for the teaching during the term of work. At the conclusion of the time selected, a second form of the same reading test which was originally given should be used with the pupils, and the papers scored as before. This will result in a final reading score for each pupil. In the case cited the final reading scores are as shown in the tenth column of Table CXIV, which should be kept by the teacher for reference during the period of the learning and be completed at its end.

TABLE CXIV. TECHNIC FOR DETERMINING OBJECTIVES IN READING (ADDITION OF FINAL READING SCORES)

PUPIL	CA	IReS	IReA	IMA	IAR	IQ	EST. FReA	EST. FReS	*FReS*
A	142	17	132	127	104	89	141	21	*20*
B	137	19	139	137	101	100	149	23	*23*
C	116	18	135	148	91	128	148	23	*24*

STEP 10. *Determining the final reading age (FReA)*

The final reading age, necessary to determine the final achievement ratio, can be found from the raw-score-age

table previously used. The ages for the three pupils cited are as shown in the eleventh column of Table CXV.

TABLE CXV. TECHNIC FOR DETERMINING OBJECTIVES IN READING (ADDITION OF FINAL READING AGES)

PUPIL	CA	IRE S	IRE A	IMA	IAR	IQ	EST. F REA	EST. F RES	FRE S	*FRE A*
A	142	17	132	127	104	89	141	21	20	*138*
B	137	19	139	137	101	100	149	23	23	*150*
C	116	18	135	148	91	128	148	23	24	*155*

STEP 11. *Determining the estimated final mental age (Est. FMA)*

The estimated final mental age is found through the same formula as that used for the other estimated ages:

$$\frac{(IQ \times \text{length of term})}{100} + \text{initial age} = \text{final age.}$$

In the case of Pupil A the substitution becomes

$$\frac{(89\ (IQ) \times 10\ (\text{months}))}{100} + 127\ (IMA) = 136\ (FMA).$$

The final mental ages found by this formula for the three pupils cited are shown in the twelfth column of Table CXVI.

STEP 12. *Determining the final achievement ratio (FAR)*

A comparison of the final achievement ratios with the corresponding initial achievement ratios will show whether the pupils have been progressing as well as they might, and how well the teacher has attained the desired objectives. These ratios are found (see page 295) by dividing the final reading ages by the corresponding final mental ages. When this was done for the three pupils here considered, the final achievement ratios were placed in the thirteenth column of Table CXVII. This table shows the completed technic for determining objectives in reading and the comparative efficiency of the teacher.

TABLE CXVI. TECHNIC FOR DETERMINING OBJECTIVES IN READING (ADDITION OF FINAL MENTAL AGES)

PUPIL	CA	IReS	IReA	IMA	IAR	IQ	EST. FReA	EST. FReS	FReS	FReA	*EST. FMA*
A	142	17	132	127	104	89	141	21	20	138	*136*
B	137	19	139	137	101	100	149	23	23	150	*147*
C	116	18	135	148	91	128	148	23	24	155	*161*

TABLE CXVII. COMPLETED TECHNIC FOR DETERMINING OBJECTIVES IN READING (ADDITION OF FINAL ACHIEVEMENT RATIOS)

PUPIL	CA	IReS	IReA	IMA	IAR	IQ	EST. FReA	EST. FReS	FReS	FReA	EST. FMA	*FAR*
A	142	17	132	127	104	89	141	21	20	138	136	*101*
B	137	19	139	137	101	100	149	23	23	150	147	*102*
C	116	18	135	148	91	128	148	23	24	155	161	*96*

As has already been indicated, the teacher who desires to do so, and who has chosen a reading test or scale which has more than two equivalent forms, can make interim checks upon the progress of the pupils in addition to the final reading objectives set up originally. Thus it would be possible to add to this reading-objectives sheet estimated objectives, ages, and scores for two, four, six, and eight months, as well as for ten, using the various periods in the formula given above instead of the ten-month period there used.

When the second form of the test which is used is not exactly equivalent in value to the first form, as in the case of the arithmetic test in the T-scale technic given below, the teacher should remember to use the method there described, which is just as applicable to age tables as to the T tables with which it is described.[1]

Determination of reasonable objectives, using the T scale. When the T scale is utilized for determining reasonable objectives in attainment for pupils, the general sequence of operations has the same purpose as when the age scale is used. By using a test with several forms it is possible to measure the points to which pupils have achieved and then, by knowing the pupils' normal rates of growth, to estimate the amounts of growth which can or might be expected during a given period.

STEP 1. *Determining the chronological age of pupils at the beginning of the growth period (ICA)*

This can be called the initial chronological age (ICA), and is found as described on page 94. In the example to be used here the chronological ages of three pupils are as shown in Table CXVIII, second column.

STEP 2. *Determining the initial arithmetic score (IArS).*

In the case of the three pupils here cited, the subject matter of the technic for determination of objectives is that

[1] See page 489.

of arithmetic, and the test used is the Woody-McCall Test in Mixed Fundamentals, of which there are four forms that are not exactly equivalent to one another. Form I was given to the pupils at the beginning of the school year, and the period selected for the attainment of the objectives was that school year, or ten months. As a result of scoring the papers of the pupils on Form I, the three pupils achieved the scores as found in the third column of Table CXIX.

TABLE CXVIII. TECHNIC FOR DETERMINING OBJECTIVES IN ARITHMETIC (T SCALE). (NAMES OF PUPILS AND CHRONOLOGICAL AGES)

PUPIL	ICA
A	12–5
B	9–6
C	13–4

STEP 3. *Determining the initial T score in arithmetic (IT_a)*

The initial T score in arithmetic is found from a table of T scores which is either published with the test or can be made according to the directions given on page 240. When T tables are not available for tests, they can easily be constructed by this means; when they are available, some time will be saved by procuring and using them.[1] In this case the T tables were procured for all forms of the test which was used, and the raw scores were transmuted directly into T scores by means of them. When this was accomplished, the T scores as derived from the T table for Form I were placed in the fourth column, as shown in Table CXX.

TABLE CXIX. TECHNIC FOR DETERMINING OBJECTIVES IN ARITHMETIC (ADDITION OF INITIAL ARITHMETIC SCORES (WOODY-MCCALL, FORM I))

PUPIL	ICA	*IArS*
A	12–5	*21*
B	9–6	*20*
C	13–4	*21*

[1] T tables especially constructed for some of the more widely used educational and intelligence tests have been constructed by Dr. McCall and his students and are available at a nominal cost through Teachers College Bureau of Publications, Columbia University, New York City.

STEP 4. *Determining the final chronological age (FCA)*

The final chronological age is found through the following simple formula: $ICA + \text{period} = FCA$.

For Pupil A, with an initial chronological age of 12–5, and ten months of learning for the attainment of the objectives, the substitution becomes

$$(12\text{–}5) + 10 = (13\text{–}3).$$

The chronological age of Pupil A at the end of the learning period will then be 13–3. When the chronological ages for the remaining pupils were found, they were included in the fifth column, as shown in Table CXXI.

TABLE CXX. TECHNIC FOR DETERMINING OBJECTIVES IN ARITHMETIC (ADDITION OF INITIAL ARITHMETIC T SCORES)

PUPIL	ICA	IArS	IT_a
A	12–5	21	*41*
B	9–6	20	*40*
C	13–4	21	*41*

TABLE CXXI. TECHNIC FOR DETERMINING OBJECTIVES IN ARITHMETIC (ADDITION OF FINAL CHRONOLOGICAL AGES)

PUPIL	ICA	IArS	IT_a	*FCA*
A	12–5	21	41	*13–3*
B	9–6	20	40	*10–4*
C	13–4	21	41	*14–2*

STEP 5. *Determining the initial B correction (IBCr)*

It will be remembered that the B correction was a figure derived from a table constructed by Dr. McCall, which, when added to or subtracted from the T score of a pupil, resulted in reducing the T score to a measure which allowed automatic comparison of the achievement, or intelligence, of a pupil with that of normal pupils of his own age. If, then, the difference between the B correction for a pupil at the beginning of the learning period and at the end of the learning period is added to his initial T score, an estimated final T score can be obtained. This step then becomes the first unit of this particular operation.

By consulting Table XLVI, p. 275, which shows the B corrections for various ages, the B correction for the initial ages can be obtained. For the three pupils cited these are found in the sixth column of Table CXXII. The B correction for Pupil A, as found in Table XLVI, is + 1 for an ICA of 12–5.

TABLE CXXII. TECHNIC FOR DETERMINING OBJECTIVES IN ARITHMETIC (ADDITION OF INITIAL B CORRECTIONS)

PUPIL	ICA	IArS	IT_a	FCA	*IBCr*
A	12–5	21	41	13–3	+1
B	9–6	20	40	10–4	+15
C	13–4	21	41	14–2	−3

STEP 6. *Determining the final B correction (FBCr)*

Since the pupil will have progressed ten months by the time of the completion of the learning period in view, the B correction for the new chronological age must be found from the final chronological ages, here included in the fifth column. The final B corrections were found and were included in Table CXXIII, as determined by consulting Table XLVI and using the final chronological ages as the guide.

TABLE CXXIII. TECHNIC FOR DETERMINING OBJECTIVES IN ARITHMETIC (ADDITION OF FINAL B CORRECTIONS)

PUPIL	ICA	IArS	IT_a	FCA	IBCr	*FBCr*
A	12–5	21	41	13–3	+1	*−3*
B	9–6	20	40	10–4	+15	*+10*
C	13–4	21	41	14–2	−3	*−6*

STEP 7. *Determining differences in initial and final B corrections*

The difference between the initial and final B corrections is found by subtracting (regarding signs) the final B correction from the initial B correction according to the formula

$$(IBCr) - (FBCr) = D.$$

The substitution for the three pupils under consideration is as follows:

For Pupil A: $(+1) - (-3) = 1 + 3 = 4(D)$.
For Pupil B: $(+15) - (+10) = 15 - 10 = 5(D)$.
For Pupil C: $(-3) - (-6) = -3 + 6 = 3(D)$.

These differences are listed in the eighth column of Table CXXIV.

TABLE CXXIV. TECHNIC FOR DETERMINING OBJECTIVES IN ARITHMETIC (ADDITION OF DIFFERENCES OF INITIAL AND FINAL B CORRECTIONS)

PUPIL	ICA	IArS	IT_a	FCA	IBCr	FBCr	*BCrD*
A	12–5	21	41	13–3	+1	−3	*4*
B	9–6	20	40	10–4	+15	+10	*5*
C	13–4	21	41	14–2	−3	−6	*3*

STEP 8. *Determining the estimated final T score* ($Est.FT_a$)

The estimated final T score is the initial T score in arithmetic added to the difference between the initial and final B corrections for each pupil, as found in Step 7.

$$IT_a + BCrD = Est.FT_a.$$

Therefore, adding the differences in the B corrections, as found in Step 7, to the initial T scores in arithmetic, as found in Step 2, will result in the estimated final T scores for each pupil, as shown in the ninth column of Table CXXV.

TABLE CXXV. TECHNIC FOR DETERMINING OBJECTIVES IN ARITHMETIC (ADDITION OF ESTIMATED FINAL T SCORES IN ARITHMETIC)

PUPIL	ICA	IArS	IT_a	FCA	IBCr	FBCr	BCrD	*Est.* FT_a
A	12–5	21	41	13–3	+1	−3	4	*45*
B	9–6	20	40	10–4	+15	+10	5	*45*
C	13–4	21	41	14–2	+3	−6	3	*44*

STEP 9. *Determining the estimated final arithmetic score (Est. FArS)*

The estimated final arithmetic raw score should be estimated for the particular form of the test which it is purposed to give to the pupils if, as in the case of the Woody-McCall arithmetic test, the forms differ at all in their values. Since it was decided that Form IV of this test was to be used for the final determination of the pupils' development, the T table for Form IV was used at this stage of the computations. In effect the process is a reversal of the process for transmuting raw scores into T scores. It consists in finding the estimated final T score in arithmetic on the raw-score-T table, and from that determining the appropriate raw score. When this step had been taken in the case here cited, the scores for the three pupils appeared as in the tenth column of Table CXXVI.

TABLE CXXVI. TECHNIC FOR DETERMINING OBJECTIVES IN ARITHMETIC (ADDITION OF ESTIMATED FINAL ARITHMETIC SCORES)

PUPIL	ICA	IArS	IT_a	FCA	IBCR	FBCR	BCRD	EST. FT_a	*EST. FArS*
A	12–5	21	41	13–3	+1	−3	4	45	*24*
B	9–6	20	40	10–4	+15	+10	5	45	*24*
C	13–4	21	41	14–2	−3	−6	3	44	*23*

STEP 10. *Determining the pupils' final raw scores in arithmetic (FArS)*

This step involves giving a test at the conclusion of the learning period determined upon (in this case ten months) and the determination of the raw scores made by the pupils on this final test. In this case Form IV of the Woody-McCall Arithmetic Test was used, and the raw scores for the three pupils were placed in the eleventh column, as shown in Table CXXVII.

STEP 11. *Determining the final T scores in arithmetic (FT_a)*

The final T scores in arithmetic are found from the T tables referred to above for the particular form of the test used

(in this case Form IV) for the determination of the final measurement. When the proper T table was selected, the raw scores were found on the table, and the appropriate T scores were placed in the twelfth column, as shown in Table CXXVIII.

STEP 12. *Determining each pupil's B score in intelligence*

Since the B scores for a pupil, for both the initial accomplishment and the final accomplishment, are of interest and help in determining his work, and since the determination of effort scores is a simple matter if the B_i score is known in addition, the B_i score of each pupil should be found. If this score has already been determined for a pupil, it need not be redetermined, because, like the intelligence quotient, it is not subject to much change. If it has not been determined and an intelligence test has already been given, the pupil's B_i score can be found by determining the T score and then adding to it, or subtracting from it, his B correction *at the time the test was given*. If none of these are possibilities, an intelligence test should be given and the B score in intelligence determined for each pupil according to the method described on page 272. When the B_i scores for the three pupils here considered had been determined, they were placed in the thirteenth column, as shown in Table CXXIX. These scores should be placed on the tabulation sheet when the determination of the arithmetic objectives is begun. They are placed in this part of the tabulation for convenience in calculation, since most of the calculations can be done mentally.

STEP 13. *Determining the initial B score in arithmetic* (IB_a)

The initial B score in arithmetic will tell the teacher what has been the past rate of learning for each of the pupils in comparison with normal pupils and, taken in connection with the B_i and the final B_a scores, will show how well each pupil has attained in terms of his own ability and promise.

TABLE CXXVII. TECHNIC FOR DETERMINING OBJECTIVES IN ARITHMETIC (ADDITION OF FINAL ARITHMETIC SCORES)

PUPIL	ICA	IArS	IT_a	FCA	IBCr	FBCr	BCrD	Est. FT_a	Est. FArS	*FArS*
A	12–5	21	41	13–3	+1	−3	4	45	24	*22*
B	9–6	20	40	10–4	+15	+10	5	45	24	*25*
C	13–4	21	41	14–2	−3	−6	3	44	23	*24*

TABLE CXXVIII. TECHNIC FOR DETERMINING OBJECTIVES IN ARITHMETIC (ADDITION OF FINAL T SCORES IN ARITHMETIC)

PUPIL	ICA	IArS	IT_a	FCA	IBCr	FBCr	BCrD	Est. FT_a	Est. FArS	FArS	FT_a
A	12–5	21	41	13–3	+1	−3	4	45	24	22	*42*
B	9–6	20	40	10–4	+15	+10	5	45	24	25	*48*
C	13–4	21	41	14–2	−3	−6	3	44	23	24	*46*

TABLE CXXIX. TECHNIC FOR DETERMINING OBJECTIVES IN ARITHMETIC (ADDITION OF B SCORES IN INTELLIGENCE)

PUPIL	ICA	IArS	IT_a	FCA	IBCr	FBCr	BCrD	Est. FT_a	Est. FArS	FArS	FT_a	B_i
A	12–5	21	41	13–3	+1	−3	4	45	24	22	42	*41*
B	9–6	20	40	10–4	+15	+10	5	45	24	25	48	*60*
C	13–4	21	41	14–2	−3	−6	3	44	23	24	46	*49*

The initial B score in arithmetic is found by adding the initial B correction (regarding signs) to the initial T score in arithmetic.[1] These scores are found in the columns headed "IT_a" and "IBCr." In the case of Pupil A, whose IT_a is 41 and whose IBCr is + 1, the calculation is

$$41 + 1 = 42\ (IB_a).$$

When the calculations were made for the other pupils, they were placed in the fourteenth column, as shown in Table CXXX. Pupil B has an IB_a of 55, and Pupil C has an IB_a of 38.

STEP 14. *Determining the final B score in arithmetic* (FB_a)

This score is determined like that in Step 13, using the data in the columns headed "FT_a" and "FBCr." In this case the final B correction is added to, or subtracted from, the final T score in arithmetic. When this step had been completed, the results were placed in the fifteenth column, as shown in Table CXXXI, p. 493.

STEP 15. *Determining the initial effort score* (IF_a)

As an evidence of the efficiency of the teaching a comparison of the initial and final effort scores would be helpful, should indicate normal effort, and should help the pupils to progress at rates normal for themselves. If the mean of the effort scores approached more nearly to 50, the teacher could feel that his teaching was approaching greater efficiency with his class.

The formula for finding the effort scores (see page 300) is

$$(B_e + 50) - B_i = F_e.$$

To find the initial effort score, therefore, the formula would be

$$(IB_a + 50) - B_i = IF_a.$$

[1] See page 282.

TABLE CXXX. TECHNIC FOR DETERMINING OBJECTIVES IN ARITHMETIC (ADDITION OF INITIAL B SCORES IN ARITHMETIC)

PUPIL	ICA	IARS	IT_a	FCA	IBCR	FBCR	BCRD	EST. FT_a	EST. FARS	FARS	FT_a	B_i	IB_a
A	12–5	21	41	13–3	+1	−3	4	45	24	22	42	41	*42*
B	9–6	20	40	10–4	+15	+10	5	45	24	25	48	60	*55*
C	13–4	21	41	14–2	−3	−6	3	44	23	24	46	49	*38*

TABLE CXXXI. TECHNIC FOR DETERMINING OBJECTIVES IN ARITHMETIC (ADDITION OF FINAL B SCORES IN ARITHMETIC)

PUPIL	ICA	IARS	IT_a	FCA	IBCR	FBCR	BCRD	EST. FT_a	EST. FARS	FARS	FT_a	B_i	IB_a	FB_a
A	12–5	21	41	13–3	+1	−3	4	45	24	22	42	41	42	*39*
B	9–6	20	40	10–4	+15	+10	5	45	24	25	48	60	55	*58*
C	13–4	21	41	14–2	−3	−6	3	44	23	24	46	49	38	*40*

Substituting in the case of Pupil A, the result becomes

$$(42\ (IB_a) + 50) - 41\ (B_i) = 51\ (IF_a).$$

The sixteenth column of Table CXXXII shows the IF_a scores for the three pupils here considered.

STEP 16. *Determining the final effort scores* (FF_a)

The final effort scores are found similarly to the initial effort scores of Step 15 by substituting the final B scores in the formula as follows:

$$(FB_a + 50) - B_i = FF_a.$$

It will be remembered that the B_i score is relatively constant when carefully determined.

When the substitutions in the foregoing formula have been made, as in the case of Step 15, the results appear as in the seventeenth column of Table CXXXIII.

STEP 17. *Determining the mean difference of the initial and final effort scores*

The mean difference of the initial and final effort scores may be easily found by direct subtraction of the initial effort scores from the final effort scores. The differences should be treated algebraically, the sum of the minus differences subtracted from the sum of the plus differences, and the final sum divided by the number of cases. The result will be the mean difference, and will show the average amount per pupil by which a teacher has helped the pupils to change their effort scores. In the case cited the differences are -3, $+3$, and $+2$, making a total difference of $+2$ $(+5-3)$, which, divided by the number of cases, gives a result of $+\frac{2}{3}$. On the average, then, the teacher has raised the effort score two thirds of one unit for each pupil. This is shown in the last column of Table CXXXIV, which shows the completed table for the T-scale technic in the determination of reasonable objectives.

TABLE CXXXII. TECHNIC FOR DETERMINING OBJECTIVES IN ARITHMETIC (ADDITION OF INITIAL EFFORT SCORES IN ARITHMETIC)

PUPIL	ICA	IARS	IT_a	FCA	IBCR	FBCR	BCRD	EST. FT_a	EST. FARS	FARS	FT_a	B_i	IB_a	FB_a	IF_a
A	12–5	21	41	13–3	+1	−3	4	45	24	22	42	41	42	39	*51*
B	9–6	20	40	10–4	+15	+10	5	45	24	25	48	60	55	58	*45*
C	13–4	21	41	14–2	−3	−6	3	44	23	24	46	49	38	40	*39*

TABLE CXXXIII. TECHNIC FOR DETERMINING OBJECTIVES IN ARITHMETIC (ADDITION OF FINAL EFFORT SCORES)

PUPIL	ICA	IARS	IT_a	FCA	I BCR	F BCR	BCRD	EST FT_a	EST. FARS	FARS	FT_a	B_i	IB_a	FB_a	IF_a	FF_a
A	12–5	21	41	13–3	+ 1	− 3	4	45	24	22	42	41	42	39	51	*48*
B	9–6	16	40	10–4	+ 15	+ 10	5	45	24	25	48	60	55	58	45	*48*
C	13–4	21	41	14–2	− 3	− 6	3	44	23	24	46	49	38	40	39	*41*

TABLE CXXXIV. TECHNIC FOR DETERMINING OBJECTIVES IN ARITHMETIC (SHOWING ADDITION OF DIFFERENCES AND THE MEAN DIFFERENCE INDICATING EFFICIENCY)

PUPIL	ICA	IAR S	IT_a	FCA	I BCR	F BCR	BCRD	EST. FT_a	EST. F ARS	FARS	FT_a	B_i	IB_a	FB_a	IF_a	FF_a	*D.*
A	12–5	21	41	13–3	+1	−3	4	45	24	22	42	41	42	39	51	48	*−3*
B	9–6	20	40	10–4	+15	+10	5	45	24	25	48	60	55	58	45	48	*+3*
C	13–4	21	41	14–2	−3	−6	3	44	23	24	46	49	38	40	39	41	*+2*
																Mean Difference	$+\frac{2}{3}$

Since the method described above for determining the teaching efficiency does not take into account the fact that the significant thing is not so much the total amount, or mean amount, which a class is raised in effort, but rather the degree to which it is helped to approach an effort rating of 50, a somewhat longer process will give a more accurate rating. This process is merely the finding of the pupils' mean initial and final effort scores. It is possible that even if the mean initial effort score were 51 and the final effort score were 55, the initial effort would be more satisfactory and would indicate more efficient teaching from the point of view of the school work as a whole than would the final effort score. Thus, if the mean of the final effort scores is nearer 50 than the mean of the initial effort scores, the teacher has been more efficient.

Use of practice tests. Practice tests, as has been stated, are now being published in several fields and are designed for continuous use by the classroom teacher for purposes of teaching. The number of fields in which these tests are available is somewhat limited, being chiefly confined to arithmetic and reading; but they can be utilized in any field where there is any large body of more or less universalized knowledge or technic and where local curricula can be adjusted for their use.

The purpose of practice tests is to present subject matter or technics in sequences which have been found desirable for learning and in a scope which it is difficult for the average teacher to equal. They are therefore extremely useful wherever standard forms of knowledge or skill obtain and where teachers are skilled enough in the use of tests to utilize them effectively.

As a general rule practice tests are so made up as to present practice materials for the use of pupils in the learning process. Interspersed among these materials are tests based upon them which are similar in nature to standard tests and which

can be treated in exactly the same way. The use of the practice materials by the pupils is designed to teach them the processes or to give them the desired units of knowledge, and the check tests are for the purpose of discovering how well the study has been undertaken.

The advantages of practice tests can be appreciated best by the teacher who uses them. They are compact and complete, covering within the range of their subject matter the necessary matters of knowledge, technic, or skill. As they are now published the cost per pupil is not great. Since they are usually made so as to be used over and over again, with the reactions of the pupils placed on tissue sheets, after the first cost has been met, the later cost is merely that of the pads of tissue for the pupils' use. Aside from such considerations of economics and curriculum, the use of practice tests within a classroom has a far more important and worthy bearing: because of their arrangement and content the pupil, as well as the teacher, has a constant check upon his own progress and, even more important, a constant check upon the causes of his failure to progress. Such a condition frees the teacher from the necessity for constant negative criticism, enables him to show the pupil how to work to better advantage, with a basis of fact which is known to the pupil as proof rather than as mere opinion, and encourages within the pupil proper habits of study, more desirable attitudes toward the subject matter and learning, and worthy incentives for progress.

The limitations of practice tests should be recognized by those who deal with them. Practice tests do not allow for local forms of subject matter or for local or provincialized forms of learning. If there are such forms within a curriculum they cannot be cared for through practice tests. This is largely the reason for the limited number of fields in which practice tests are available, although (as will be shown later)

the testing of these localized forms of subject matter can be cared for in other ways. As has been implied, the practice tests are limited in the number of fields in which they can be effectively used as tests. However, the technics which make practice tests valuable as aids in teaching can be separated from the technics of testing as such, and consequently many publications are appearing on the market which have the form of practice tests but which are called, instead, practice materials, and which are of great help in the teaching of the technical aspects of geography, science, history, and the like. Such materials, valuable as they are and increasingly valuable as they are sure to become, are not tests and (at the present time at least) cannot be treated as tests.

Use of classroom tests. Classroom tests, as distinguished from standard tests,[1] are unstandardized. They are made by the teacher, and the results from them can be used only within the classroom where they are given. The purpose of making classroom tests is to supplement the use of standard tests largely in the fields of subject matter where standard tests are as yet not well developed, where there is considerable difference in the content of subject matter from school to school or from system to system, and where additional measures are desired.

Classroom tests are designed to correct some of the difficulties in judgment on the part of teachers and objectivity in scoring, which were outlined in Chapter XVII, and are particularly designed to replace (at least in part) the traditional forms of teacher-made tests and examinations. To increase the range and objectivity of local testing by the teacher, classroom tests adapt, so far as they can, the technics accepted in the form of standard tests. The range of the test is increased over that of the traditional forms through an increase in the number of items included within it, the time

[1] See page 40.

taken by the test in relation to the number of items is reduced through requiring only short and definite answers, and the objectivity of the testing is increased through eliminating, so far as possible, the subjective judgments of the teacher in the scoring.

Various forms of the tests have been developed in order to measure a wider range of abilities, or development, than was possible under the more traditional forms of testing by the teacher, and the technics of making and giving them have become fairly well established. Although many different forms of classroom tests have been proposed, those which have seemed to be most successful in supplementing the standard-test forms are the following:

1. *The true-false test.* In this form of test a number of items are presented to pupils, some of which are true and others of which are not true. The pupils are asked to indicate which are true and which are not. Range is increased through the number of items in the test, and objectivity is increased through an impersonal form of scoring. The test is adapted to a wide variety of subject matter and is increasingly useful in determining the reaction of pupils to situations like those which obtain in their daily living. There are many objections to the use of the true-false test, chiefly concerned with the presentation of false items to pupils and the methods of scoring the tests. In spite of these objections, which are undoubtedly valid in many instances, proper precautions and the use of accepted technics and subject matter make the tests of great value to teachers.

2. *The judgment test.* This test is designed to parallel in its measurement the reasoning abilities which it was formerly attempted to measure through the traditional essay form of examination. A group of true statements is presented to the pupils, and they are asked to tell why the statements are true. The answers thus all begin with the word "Because."

This test is somewhat more difficult to score than most of the other forms of classroom tests, and is more limited in range than many of them. Its great value lies in the results which it brings, since a teacher is able to locate with considerable accuracy the causes of his pupils' difficulties in understanding. The scoring is less objective than that of the true-false test and other types, but technics have been developed whereby the pupils can be protected and the teacher can be fair. The test is adapted to as wide a range of subject matter as is the true-false test and is particularly of value when quality of judgment is a paramount factor.

3. *The selection test.* This test appears in a great many forms, variously called the multiple-choice test, the two-column selection, the reorganization test, and the like. In each form of this test the object in the testing is to present to pupils series of facts out of their natural order, or sequence, with the idea of having the materials properly reassembled or selected. The test is therefore one of intellectual discrimination, with many values in teaching. It is, as a rule, an easier and more enjoyable form of test to pupils than others; therefore it is more eagerly welcomed by them and is attacked with zest. The methods of scoring the test are usually simple and objective, and the results are quickly found. It is adaptable to a wide variety of subject matter and can therefore be used with great frequency and effect.

4. *The association test.* This test has some of the characteristics of the judgment test, but it is considerably longer in range, is somewhat more limited in its use, and is somewhat easier to score. It presents to pupils a number of words or phrases of technical or semi-technical character and asks for a short simple statement about each of them. The pupils respond by stating what is to them their most important association with the facts or ideas, and the teacher is thereby

informed of the quality of their associations with these. This test is more subject than others to the abuses of formalism, but when used with care, it has many values for quick and informative testing.

5. *The completion test.* Adapted directly in both form and technic from certain types of standard tests, the completion test may easily become one of the more difficult of the classroom forms. It presents to pupils a series of statements which are true, but in which certain essential ideas have been concealed by blank spaces. It tests not only knowledge but also literacy, and thus combines some of the more valuable purposes of the traditional forms of testing, giving at the same time a greater objectivity in the judgment of answers, together with the possibility of wider range. The test is widely adaptable throughout school work and has proved to be of great value.

In the classroom these tests can be combined with each other in varying proportions or even with the traditional forms, and thus a variety of abilities or results of learning can be measured at one sitting. In general, with the observation of reasonable precautions in making and giving the tests, they are not difficult to construct or administer, although they do consume more time in the making than do the essay types. Their great advantage in this connection, however, is that the time which is spent in making them is more than compensated for by the saving of time in their correction and by the greater usefulness of the results.

Many of the desirable results of standard testing can be achieved through classroom testing. They cannot be used for purposes of comparison of class with class or of school with school, and can be used only indirectly in the classification of pupils. From the results of classroom tests, however, it is possible to determine relative question difficulty, to diagnose

many of the difficulties of pupils, to determine both achievement ratios and effort scores, to combine accurately the results of several tests into accurate composites, and to use the results of the tests (see Chapter XVII) for marking and grading. The only real use which can be made of classroom tests by the teacher in terms of the classification of pupils is to reenforce his judgment with respect to the relative worth of pupils for promotion. This, it will be remembered, was one of the items which was included in the classification of pupils,[1] and the only one in which the teacher's judgment was brought into account.[2]

Classification of pupils within classes. For purposes of instruction it is frequently desirable to separate the pupils into groups. These groupings may be in terms of ability, achievement in a given subject, degree of effort expended, or the like. Whatever purpose the teacher may have in mind, there are two general forms of classification groupings which may be made. On the one hand, the pupils may be divided into several complementary groupings. In this form of grouping the pupils in any one group differ from those in any other group, and the groups are stratified throughout the class. If there are two complementary groups, one consists of the better half of the class, and the other of the lower half; if there are three groups, one consists of the upper third of the class, one of the middle third, and the remainder make up the lowest third. On the other hand, there may be equivalent groupings of the pupils, whereby each of the groups is equivalent to any of the others in terms of the basis on which they have been separated.

1. *Complementary groupings.* When pupils are to be divided into complementary groups the technic is as follows:

[1] See page 320.

[2] For a more complete and detailed account of the making, giving, and using of classroom tests, see Charles Russell's *Classroom Tests* (Ginn and Company, Boston, 1926).

STEP 1. *Placing the pupils in order*

When the particular basis of division of the pupils has been selected, such as achievement in reading, effort in reading, general ability, achievement in arithmetic, and the like, the first step is to list them in order according to the scores which they have made. If it is supposed that the fifth-grade pupils shown in Fig. 94 are, on the basis of their achievement in arithmetic, to be divided into three complementary groups, the order of their scores would appear as shown in Table CXXXV. In this case the T_a scores are being used. The raw scores could have been used just as well. B scores should not be used, because what is desired in ranking is not a relationship to normal for the ranking, but rather a relationship to the other pupils within the class. Either T_a scores, or raw scores in arithmetic, or age scores in arithmetic could be used. The first column shows the names of the pupils, indicated by the same letters as were used in Fig. 94; the second column shows the T scores in arithmetic.

TABLE CXXXV. DIVISION OF PUPILS INTO THREE COMPLEMENTARY GROUPS

GROUP 1		GROUP 2		GROUP 3	
Pupils	T_a	Pupils	T_a	Pupils	T_a
H	70	Q	48	F	41
R	59	D	46	M	41
DD	56	E	46	O	41
X	54	G	46	V	41
Y	54	S	46	Z	41
FF	51	KK	46	AA	41
HH	51	T	44	BB	41
C	50	W	44	II	41
L	50	K	42	JJ	38
U	50	P	42	A	34
GG	50	CC	42	N	34
I	48	B	41	EE	32
J	48				

STEP 2. *Dividing the number of pupils by the number of groups desired and separating the groups accordingly*

If there were to be two groups, and if there were thirty pupils, each group would include fifteen pupils ($\frac{30}{2}$); one group would include the first, or better, fifteen pupils, whereas the second group would include the lower fifteen.

TABLE CXXXVI. ASSIGNMENT OF PUPILS TO TWO EQUIVALENT GROUPS

PUPIL	T_a	GROUP	PUPIL	T_a	GROUP
H	70	1	T	44	1
R	59	2	W	44	1
DD	56	2	K	42	2
X	54	1	P	42	2
Y	54	1	CC	42	1
FF	51	2	B	41	1
HH	51	2	F	41	2
C	50	1	M	41	2
L	50	1	O	41	1
U	50	2	V	41	1
GG	50	2	Z	41	2
I	48	1	AA	41	2
J	48	1	BB	41	1
Q	48	2	II	41	1
D	46	2	JJ	38	2
E	46	1	A	34	2
G	46	1	N	34	1
S	46	2	EE	32	1
KK	46	2			

In this case there were thirty-seven pupils and three groups. Thirty-seven should be divided by three, giving two groups of twelve pupils each and one group of thirteen. It is probable in this case that the best group will need the least help from the teacher. This group can therefore include the thirteen pupils, and the other two groups twelve pupils each. The columns in Table CXXXV indicate Group 1, the best group as containing the first thirteen pupils, and show how the remaining two groups are selected. If the pupils are divided

for purposes of special instruction, Pupil Q should be placed in Group 1 and Pupil B in Group 3, because of their scores.

2. *Equivalent groups.* When pupils are to be divided into equivalent groups the technic is as follows:

STEP 1. *Placing the pupils in rank order*

In separating classes into equivalent groupings it is desirable to have the groups as nearly equivalent as possible. This necessitates some care in the selection of the pupils, although the technic of selecting them is the same regardless of the number of equivalent groups for any one class.

Let it be supposed that the class in arithmetic referred to above is to be divided into two equivalent groups. In this case the pupils should first be placed in order, as shown in Table CXXXVI, second column.

STEP 2. *Assigning the groups according to the sequence*

The sequence of assigning the pupils to two groups is as follows:

GROUP NUMBERS: 1 2 2 1 1 2 2 1 1 2 2 etc.
PUPIL NUMBERS: 1 2 3 4 5 6 7 8 9 10 11 etc.

This means that Pupil 1 is assigned to Group 1, Pupil 2 to Group 2, Pupil 3 to Group 2, Pupil 4 to Group 1, and so on. This sequence has been followed in the third column of Table CXXXVI. Table CXXXVII shows the two groups separated.

If the class were to be divided into three equivalent groups, the sequence of assignment would be as follows:

GROUP NUMBERS: 1 2 3 3 2 1 1 2 3 3 2 1 1 etc.
PUPIL NUMBERS: 1 2 3 4 5 6 7 8 9 10 11 12 13 etc.

In this case the first three pupils would be assigned to the three groups in order, the fourth pupil to the third group, the fifth to the second, the sixth and seventh to the first, and so on.

TABLE CXXXVII. PUPILS OF TABLE CXXXVI DIVIDED INTO TWO EQUIVALENT GROUPS

GROUP 1	T_a SCORE	GROUP 2	T_a SCORE
H	70	R	59
X	54	DD	56
Y	54	FF	51
C	50	HH	51
L	50	U	50
I	48	GG	50
J	48	Q	48
E	46	D	46
G	46	S	46
T	44	KK	46
W	44	K	42
CC	42	P	42
B	41	F	41
O	41	M	41
V	41	Z	41
BB	41	AA	41
II	41	JJ	38
N	34	A	34
EE	32		

Table CXXXVIII shows how this series of assignments would be made; Table CXXXIX shows the separated groups.

Conclusion. Standard tests are unquestionably one of the more important additions of the present to the field of education. They are capable of use for the direction of pupils' work, for the stimulation of better teaching, for the reorganization of classes, grades, and schools, for the development of more adequate curricula, and for the better understanding of the educational process. They should be one of the chief tools of the teacher, the supervisor, and the superintendent.

With all their values, however, standard tests can become undesirable if they are misused or misinterpreted. One misuse is to believe that there is any intrinsic value in merely giving the tests. They are not valuable in this way and cannot replace the teaching of the curriculum: they can only be

TABLE CXXXVIII. ASSIGNMENTS OF PUPILS TO THREE EQUIVALENT GROUPS

PUPIL	T_a SCORE	GROUP	PUPIL	T_a SCORE	GROUP
H	70	1	T	44	2
R	59	2	W	44	3
DD	56	3	K	42	3
X	54	3	P	42	2
Y	54	2	CC	42	1
FF	51	1	B	41	1
HH	51	1	F	41	2
C	50	2	M	41	3
L	50	3	O	41	3
U	50	3	V	41	2
GG	50	2	Z	41	1
I	48	1	AA	41	1
J	48	1	BB	41	2
Q	48	2	II	41	3
D	46	3	JJ	38	3
E	46	3	A	34	2
G	46	2	N	34	1
S	46	1	EE	42	1
KK	46	1			

TABLE CXXXIX. DIVISION OF PUPILS INTO THREE EQUIVALENT GROUPS

GROUP 1	SCORE T_a	GROUP 2	SCORE T_a	GROUP 3	SCORE T_a
H	70	R	59	DD	56
FF	51	Y	54	X	54
HH	51	C	50	L	50
I	48	GG	50	U	50
J	48	Q	48	D	46
S	46	G	46	E	46
KK	46	T	44	W	44
CC	42	P	42	K	42
B	41	F	41	M	41
Z	41	V	41	O	41
AA	41	BB	41	II	41
N	34	A	34	JJ	38
EE	32				

of true value if they serve to readjust a pupil so that his learning can have more effect, or if they serve to readjust the teaching so that a pupil can have a better reaction. Another misuse results from the misinterpretation, or lack of interpretation, of the results of the tests. The teacher who has given the tests carefully but is unwilling to accept the results can receive but little benefit, and the pupils have derived no benefit from the testing.

The true values of standard tests can only come when the teacher uses them with intellectual humility, when he is aware of their high potential values, when he is determined to realize those values to the utmost, and when he appreciates the fact that the ultimate purpose of testing is to provide for some desirable change in pupils.

Chapter summary. Among the more important technics for the use of standard tests are those directly connected with the teaching process. In these technics the testing is incidental but necessary to the teaching, and the purpose is primarily to aid rather than to measure learning.

One technic is that of the determination of reasonable objectives for learning. It can be accomplished either in the age scale or in the T scale. In both cases the object is to give a test, to transpose the scores into initial age or T scores, to calculate the amount of growth which should take place in a given length of time (taking into consideration the pupils' rates of learning), and to transpose these estimated amounts of growth into test scores in a second form of the test originally given. This process determines the objectives of normal growth, and a comparison of these objectives with the actual accomplishment of the pupils will give a measure of teaching efficiency.

A second technic consists in the use of practice tests. These tests, composed of practice materials and standard tests, allow a teacher to give aid to a pupil when he needs it most,

permit a constant check upon the pupil's development, and show the pupil his errors and needs without the intervention of the teacher. They are extremely valuable as devices to aid teaching, but are available only in a few fields where there is a considerable body of universalized knowledge or technic.

A third technic which can be used to supplement standard tests is the construction and use of classroom tests. These are test forms constructed after the manner of standard tests, but are made by the teacher and are unstandardized. They are therefore of little use for comparative purposes or for absolute ratings. They can, however, be used to supplement the teacher's judgment in making estimates of worth for promotion, and therefore are of value in connection with the use of tests for reclassifying pupils.

A fourth technic is that of the classification of pupils within a classroom on the basis of ability, achievement, or effort. Pupils may be classified either in complementary groups, where each group differs from the others in average achievement, ability, or effort, or in equivalent groups, where each group is as nearly the equivalent of the others as may be.

SELECTED BIBLIOGRAPHY

BUCKINGHAM, B. R. Research for Teachers, pp. 133–198. Silver, Burdett and Company, New York, 1926.

FENTON, N., and WORCESTER, D. A. An Introduction to Educational Measurements, pp. 51–70. Ginn and Company, Boston, 1928.

MCCALL, W. A. How to Measure in Education, pp. 67–168. The Macmillan Company, New York, 1922.

MORT, P. R. The Individual Pupil. American Book Company, New York, 1928.

WILSON, G. M., and HOKE, K. J. How to Measure (Revised Edition), pp. 356–387. The Macmillan Company, New York, 1928.

INDEX

Ability, meaning and importance of, 248. *See also* Measures of ability
Ability measures. *See* Measures of ability
Achievement, significance of class, 410
Achievement measures. *See* Measures of achievement
Achievement quotient, determination of, 280; formula for, 281; interpretation of, 282
Achievement ratio, determination of, 296; formula for, 297; interpretation of, 298; use of, in classification, 357, 366
Achievement tests, used in classification, 308, 353
Actual classification, by T scale, 346; by age scale, 359; by age-grade scale, 366
Age, determination of chronological, 94; determination of mental, 260
Age-grade reclassification table, 360
Age-grade scale used in classification, 361
Age-grade tables, construction of, from both age and grade norms, 217; from age norms only, 224; from grade norms only, 235
Age norms, 112; for intelligence test, 259
Age scale, foreshadowed, 38; definition of, 202; measures of achievement in, 212; measures of ability in, 258; measures of effort in, 295; used in classification, 354; determination of potential ability for classes in, 411; determination of class achievement in, 418; determination of class effort in, 424; determination of objectives in, 476
Age status for promotion used in classification, 355
Age table constructed from age norms, 212
Army Intelligence Tests, 40
Association test, 500
Axioms of measurement, 44

B correction, determination of, 276
B scale, derivation of, 273
B score in education, derivation of, 282; interpretation of, 284
B score in intelligence, determination of, 272; interpretation of, 277
B score for promotion, determination of, 287; interpretation of, 289; used in classification, 333; formula for, 335
Base line, drawing of, 157
Behavioristic psychology, influence of, 24
Binet-Simon Scale for the Measurement of Intelligence, 38
Buckingham, Dr. B. R., 37
Burgess Tests of Silent Reading, 42

Chaplin, Dr. Hugh, 408
Chivalric measurement, 17
Chronological age, determination of, 94; used in classification, 314, 353, 362
Class achievement, significance of, 410; determination of, in age scale, 417; determination of, in grade scale, 418; determination of, in T scale, 420
Class effort, determination of, in age scale, 424; in T scale, 425
Classification data sheet, 312; useful information from, 349
Classification of pupils, within a school, 304; problem of, 304; proc-

ess of, by T scale, 306; process of, by age scale, 352; process of, by age and grade scales, 361; within a class, 502
Classifications, actual, by T scale, 346; by age scale, 360; by age and grade scales, 366
Classifications, statistical, by T scale, 341; by age scale, 359; by age and grade scales, 366
Classroom tests, development of, 40; characteristics of, 58; described, 498
Complementary groupings, 502
Completion test, 501
Conduct of testing period, 124
Construction of tests: identification of pupil, 93; identification of test, 97; final scores, 99; directions for giving, 100; preliminary practice exercise, 103; body of test, 105; answers, 105; directions to examiner, 107; norms for test, 110; scoring and graph sheet, 115; report to author, 115
Courtis, Dr. S. A., 37, 38
Courtis Standard Practice Tests in Arithmetic, 42
Crude mode, 168; uses of, 198
Curves, normal, 189; plus-skewed, 190; minus-skewed, 190; multimodal, 190

Diagnosis, of individuals, 368; of classes, 410; of schools, 430
Diagnosis of difficulties in learning. *See* Preliminary diagnosis of difficulties in learning, Secondary diagnosis of difficulties in learning
Diagnosis of individuals, range of, 369; as to physical difficulties, 370; as to mental difficulties, 372; as to learning difficulties, 373
Diagnosis of mental difficulties, 372
Diagnosis of physical difficulties, 370
Dictation of tests, 119
Directions to examiner, 107
Distributions, order, 146; rank, 146; frequency, 148
Divisors for finding step intervals, 151

Educational age, used in classification, 354, 362
Educational diagnosis, of individuals, 368; of classes, 410; of schools, 430
Educational objectives. *See* Objectives
Effort, relation of, to achievement, 292; importance of, in classroom, 292; essentials for measurement of, 294
Effort measures. *See* Measures of effort
Equivalent groupings, 505
Essay examinations, unreliability of, 29; restricted range of, 30

F score: meaning, 298; derivation, 299; formula, 300; interpretation, 301; used in classification, 338
Faculty psychology, influence of, 19
Fisher's *Scale Book*, 33
Fixed rating standard, establishment of, 457
Formulas: step intervals, 150; mental age, 264; intelligence quotient, 271; B score, 276; achievement quotient, 281; achievement ratio, 297; F score, 300; B score for promotion, 335
Frequency distribution, 148; uses of, 189
Frequency surface, construction of, from frequency distribution, 156; construction of, from raw scores, 159; uses of, 189; interpretation of, 191
Future rate of learning, measures of, in age scale, 285; measures of, in T scale, 287

G scores, transmutation of raw scores into, 203; preparation and use of transmutation sheet when using, 204
Gates, Dr. Arthur I., 385
Goddard, Dr. H. H., 39
Grade norms, 112; determination of, 203
Grade scale, definition of, 202; measures in, of achievement, 203;

measures in, of ability, 255; used in classification, 360; determination of potential ability for classes in, 413; determination of class achievement in, 418
Grade score in education, used in classification, 319
Grade score in intelligence, derivation of, from grade norms, 256; derivation of, from age norms, 257; derivation of, from B_i scores, 258
Grade score for promotion, used in classification, 326, 364
Grade table, construction of, from grade norms, 203
Graph sheet, use of, in frequency surface, 157
Greek measurement, 15
Gregory, Dr. C. A., 56

Hillegas, Dr. M. B., 37

Ideals, measurement of, 73
Intelligence quotient, determination of, 269; formula for, 271; interpretation of, 271
Intelligence test, Army Alpha, 40; Army Beta, 40; Otis, 40; Haggerty, 40; National, 40; individual, 77; group, 77; Pintner-Cunningham, 82; precautions in selecting, 251; precautions in giving, 252; precautions in scoring, 254; used in classification, 306

Judgment test, 499

Knowledge, measurement of, as related to skills, 66; measurement of, as related to reasoning, 68

Letter grades, establishment of fixed rating standard for, 457; technic of transmutation of standard test scores to, 459; systems of, based on normal curves, 460; systems of, based on minus-skewed curves, 461
Location of difficulties in effort (effort scores high): overstimulation, 398; lack of recognition of the limitations of ability, 399; unequal educational development, 399
Location of difficulties in effort (effort scores low): lack of recognition of ability of pupil, 396; lack of challenge in the work, 396; lack of adaptation to the work, 397; improper classification, 398
Location of difficulties in learning (previous rate faster than potential rate): improper classification, 394; undue pressures, 394; exceptional effort, 395; lack of uniform development, 395
Location of difficulties in learning (previous rate slower than potential rate): improper classification, 392; difficulties in subject matter, 392; lack of effort, 393; lack of uniform growth, 393

McCall, Dr. W. A., 56; theses of measurement of, 44; T scale of, 202
Mass measures. *See* Measures of mass
Mean, determination of, by method of summing of scores, 181; determination of, by guessed mean method, 182; uses of, 198
Meaning of scores, in spelling, 152; in composition, 153
Measure of future rate, in age scale, 285; in T scale, 287
Measure of potential rate, in age scale, 269; in T scale, 272
Measure of previous rate, in age scale, 280; in T scale, 282
Measurement, relations of, to knowledge, 3; relations of, to scientific progress, 5; relations of, to transportation, 7; relations of, to communication, 9; relations of, to food supply, 10; in education, 12; primitive, 14; in Greece, 15; in chivalry, 17; of intelligence, 38, 75; of skills, 59; of knowledge, 63; of ideals, 73
Measures of ability, in grade scale, 255; in age scale, 258; in T scale, 260

Measures of achievement, 145; in grade scale, 203; in age scale, 212; in T scale, 236
Measures of effort, 145; in age scale, 295; in T scale, 298
Measures of mass, 145; order distribution, 146; rank distribution, 146; step list for rank distribution, 147; frequency distribution, 148; step list for frequency distribution, 150; frequency surface, 156; step list for frequency surface, 156; uses of, 187
Measures of points, 144; crude mode, 168; median, 169; first quartile point, 173; third quartile point, 175; special considerations in finding median and quartile, 178; mean, 181; interpretation of, 198
Measures of rate, 145
Median, assumptions in finding, 169; steps in finding, 170; special considerations in finding, 178; uses of, 198
Mental age, meaning of, 258; derivation of, 260; keeping, up to date, 263; formula for, 264; used in classification, 353
Minus-skewed curve, 190
Morrison-McCall Spelling Scale, 67
Mort, Dr. Paul R., 401. *See also* Pupil classification card
Multimodal curve, 190

Nassau County Revision of the Hillegas Scale of English Composition, 42
National Intelligence Test (N. I. T.), 40; age norms for, 259; T–G table for, 262
Normal curve, 189
Norms, development of, 37; a characteristic of standard tests, 58; defined and illustrated, 110. *See also* Age norms, Grade norms

Objectives, determination of, in age scale, 476; in T scale, 484
Objectives of education, changes in, 18
Objectivity, 41, 56
Order distribution, determination of, 146; uses of, 187
Otis Tests of Intelligence, 40

Percentile grades, transmutation of standard test scores to, 467
Plans for testing, 132
Plus-skewed curve, 190
Point measures. *See* Measures of points
Potential class ability, in age scale, 411; in grade scale, 413; in T scale, 414
Potential rate of learning, measures of, in age scale, 269; measures of, in T scale, 272
Practice exercises, 103
Practice tests, described, 85; use of, 496
Preliminary diagnosis of difficulties in learning: determination of status in subject achievement, 374; determination of status in rate of learning, 380; determination of status in effort of learning, 382
Preparations for giving tests, 122
Previous rate of learning, measures of, in age scale, 280; measure of, in T scale, 282
Primitive measurement, 14
Profile chart, 400
Progress sheet for testing program, 139
Promotion age, used in classification, 356, 364
Promotion quotient, determination of, 285; interpretation of, 287; used in classification, 358, 364
Psychology, influence of faculty type of, 19; influence of behavioristic type of, 24
Pupil classification card (by Dr. Mort), 402–404, 442

Quartile points, first, 173; third, 175; special considerations in finding, 178; uses of, 198
Quotients, intelligence, 269; achievement, 280; promotion, 285; versus ratios, 295

Rank distribution, determination of, 146; uses of, 187
Rate in learning, significance of, 267; measures of potential, 268; measures of previous, 278; measures of, future, 284
Rate measures. *See* Measures of rate
Rating of pupils: difficulties, 453; relation to standard tests, 455; establishment of fixed rating standard, 457; technic of transmuting standard test scores to letter grades, 459; technic of transmuting standard test scores to percentile grades, 467
Ratios, versus quotients, 295; formula for, 297
Reclassification data sheet, 313
Results of testing, assembly of, 131
Rice, Dr. J. M., 33, 34, 35

Scale, definition of, 35; meaning of, 201
Scale Book, the Reverend George Fisher's, 33
School measures, traditional, 29
School surveys, 438
Scores, in spelling, 152; in composition, 153
Scoring of tests, 128
Secondary diagnosis of difficulties in learning. *See* Location of difficulties in effort, Location of difficulties in learning
Selection test, 500
Skills, measurement of, 59; related to quality, 60; related to time, 61; related to difficulty, 62
Standard tests, versus classroom tests, 56; characteristics of, 56; classification of, 59; construction of, 93; ways of giving, 119; scoring of, 128; assembly of results from, 131; relation of, to schools, 430; relation of, to the school survey, 438; relation of, to pupils in a school survey, 439; relation of, to teaching in a school survey, 443; relation of, to teachers in a school survey, 446; relation of, to curriculum, 448; relation of, to rating of pupils, 455
Standards, 37
Stanford Revision of the Binet-Simon Scale for the Measurement of Intelligence, 39
Statistical classification, by the T scale, 346; by the age scale, 359; by the age-grade scale, 366
Statistical methods, development of, 35
Step list for testing, 137
Stone, Dr. C. W., 36
Survey tests, 88

T scale, foreshadowed, 38; definition of, 202; measures in achievement, 236; derivation of, 238; meaning of, 238; measures in ability, 260; measures in effort, 299; used in classification, 306; determination of potential class ability in the, 414; determination of class achievement in the, 420; determination of class effort in the, 425; determination of objectives by means of, 484
T score in education, used in classification, 318
T table, construction of, when grade norms only are available, 240; construction of when age norms only are available, 245
Tally sheet for frequency distribution, 153
Teacher's estimate of pupils' worthiness for promotion, used in classification, 320; procedure for finding, 324; used in age-scale classification, 363
Terman, Dr. L. M., 39
Test, definition of, 35; meaning of, 201
Testing period, preparation for, 122; conduct of, 124
Testing program, plans for, 131; for a school, 135; step list for, 137; progress sheet for, 139
Tests, development of classroom, 40; practice-type standard, 85; surveys and batteries of standard, 88; construction of standard, 93; timing

standard, 125; scoring standard, 128; classroom, described, 498
Thorn Reading Tests, 42
Thorndike, Dr. E. L., 34, 35; development of handwriting, scale of, 36
Thorndike-McCall Reading Scale, 42
Thorndike Word Book, 49
Timing of tests, 125
Trabue, Dr. M. R., 33, 460
Traditional school measures, 29
Transmutation of raw scores, to grade scores, 203; to age scores, 212; to age-grade scores from age and grade norms, 217; to age-grade scores from age norms only, 224; to age-grade scores from grade norms only, 235; to T scores when grade norms only are available, 240; to T scores when age norms only are given, 245
Transmutation sheet, use of, in finding raw-score-grade table, 204; for deriving grade scores, 256; for deriving percentile grades, 467
Transmutation of standard test scores, to letter grades, 459; to percentile grades, 467
True-false test, 41, 499

Wood, Dr. Thomas D., 371, 401, 405
Woody Arithmetic Tests, 42
Woody-McCall Arithmetic Tests, 42

OCT 1 7 1990

LB1131.R87
Standard tests
WITHDRAWN
3 0629 0071411 1

LB
1131
R87